DRUG INFORMATION ASSOCIATION

FDA/DIA SCIENTIFIC WORKSHOP ON FOLLOW-ON PROTEIN PHARMACEUTICALS

PLENARY SESSION

Tuesday, February 15, 2005

8:05 a.m.

Marriott Crystal Gateway
1700 Jefferson Davis Highway
Arlington, Virginia

C O N T E N T S

P R O C E E D I N G S

DR. JONECKIS: Good morning, ladies and gentlemen. This is the second day of the follow-on protein pharmaceuticals, a joint DIA/FDA sponsored workshop.

We have another busy day, following up on yesterday's activities. You will be happy to know the moderators from the various groups worked late into the night, burning the midnight oil, trying to summarize the results of the breakouts, at least the salient points, and you will hear from those this morning.

If we can just a take a minute and look through the day's agenda.

We have the reports of three breakouts today; that is, the analytical characterization, the biological clinical, and the pharm/tox characterization and studies.

The clinical pharmacology studies will be presented tomorrow at the results of the breakouts that will be conducted this afternoon.

We will then move into two plenary

sessions, the first one being approaches to immunogenicity studies, and then approaches to clinical safety and efficacy studies.

Amy Rosenberg will be chairing that session, and we will have two speakers, both Robin Thorpe and, I'm going to butcher this name, I apologize, Huub Schellekens.

The approaches to clinical safety and efficacy study will be chaired by Dr. David Orloff, then we'll have speaker Jay Siegel, as well as Carol Ben-Maimon.

We will get the pre-lunch announcements, if we have any.

I think without further ado, what we would like to do then is start with the results of the first breakout session on analytical characterization.

Again, as I indicated, the moderators worked late into the night yesterday to try to summarize the results of those sessions, and that is what you will hear now.

As you probably have noticed and as I

believe we announced, there were transcripts that were made of the various meetings and they will be available in approximately 30 days.

The mechanism for making those available is not clear at present, for several reasons, but, hopefully, tomorrow or, if not, sometime in the future, there will be an announcement as to how those transcripts can be obtained.

I think without further ado, we can go ahead start.

First, analytical clinical presentation will be presented by Dr. Andrew Chang, one of the moderators from FDA.

DR. CHANG: Good morning. My name is Andrew Chang. Up to this week, I am still acting Deputy Director for the Division of Hematology in CBER/FDA.

I am going to give you a report on issues and consensus. Actually, you will find out we haven't gotten many consensus from Session A, entitled "Physical Chemical Characterizations and Impurities."

Now, before I do that, on behalf of the five moderators for the Session A, I would like to thank those of you that participated at our

session. We did have a very active discussion at our session, and we had a full house yesterday for our session.

I would like to thank you for your participation.

Now, we used a format that I don't know whether it is unique to our session or not. What we did is that we asked the industry moderators to present points and counterpoints to facilitate the discussion.

Also, we took each question sequentially. So we totally had three questions for discussion.

Now, we have three FDA moderators and two industry moderators for this session. FDA moderators are Barry Cherney, myself, Steve Moore. We have Charlie Di Ligerti from the Barr Laboratory, and, also, Reed Harris from Genentech.

The first question is which product attributes should be evaluated. Again, to

facilitate the discussion, Reed Harris had the following two questions and one point. What is known about molecular characteristics that mediate bioavailability, potency, safety, including immunogenicity?

Now, my intent here is really to try to recap what we had discussed, with very minimum elaboration by myself. So I very much just want to present to you what was done yesterday.

What is known about the rules of the degradation? That could be the effects of the container and/or storage conditions.

Process-related impurities, such as host cells, proteins, leachable components that could be considered as critical attributes? Of course, that would be a case-by-case situation.

Now, Charlie Di Ligerti offered some counterpoints after Reed Harris' presentation.

Perform four physical chemical characterizations using all available and relevant comparative analytical tools. Perform redundant

measurements of each aspect of a structure and impurity with multiple orthogonal methods. The follow-on industry should address identity, purity and potency. Analytical results collectively provide highly sensitive and selective fingerprint of a product.

Then we opened the floor for discussion, and I'm going to present you some of the points that we captured, identified from the audience and, also, moderators.

There is relatively good consensus that all relevant parameters should be evaluated and one comment indicated that there is no need to test it to infinity.

Historical database needed in order to have a meaningful comparability studies and in order to identify critical quality attributes.

Value of literature. The counterpoint on this is that you are able to learn a lot of information from the public literature.

Based on signs and the product clause, and there is a good consensus on that. One example

provided is that bacteria expression systems, such as expression in E. coli, and you don't need to do extensive glycosylation assays. The host does not offer the mechanism for that.

Perceived clinical issues, safety and efficacy. All properties may be relevant for safety, and some emphasis on the discussion of the quality attributes to the safety.

For example, what are the molecular attributes that you know that relate to the safety, such as immunogenicity.

So the counterpoint that all properties may be relevant for safety.

Discussion regarding what relevant means. There was some discussion on the relevant means and some clarification from Charlie that he meant relevant analytical technologies. So there was some discussion there.

Orthogonal approach needed, and there is good consensus there. You need to use the different analytical technology with different principles to measure some of the same quality

attributes. Even then, you may not find what you are not looking for.

So then we moved to question two. We gave 15 minutes for discussion for question one, 30 minutes for question two, and 40 minutes for question three.

Question two, what are the capabilities and limitations of the available analytical tools to evaluate those identified product attributes?

Now, we used the same format, that Reed Harris offered the following points.

Limits, such as the length of the proteins, modification, size, number of polypeptides dictate the limitations or capability of the analytical method.

Single modification type at multiple sites, some example was used at the plenary session in the morning yesterday, such as glycosylation, versus variable modifications at one site.

High order structure methods, and there was some discussion of what is available technology that can measure the high order structure.

Some examples provided, like deamidation, that probably were missed by a normal analytical technology used.

Glycosylation, such as N-linked or O-linked, the site occupation, terminal groups, and those are important and sometimes that was limited to analysis if the product is not very pure.

The counterpoint from Charlie, that comparative characterization is both possible and routine for most protein products and provides a foundation for supporting product changes and comparisons.

Similar logic and criteria should be used for comparisons between products from different manufacturers. Complete elucidation of covalent structure, sensitivity measures for comparing higher order structure, such as fingerprint; sensitivity measures for measuring impurities.

For question two, the discussion actually centered on limitations at our session.

Ability to detect clinically relevant properties. Again, that was also touched upon for

question one, the criteria that you should detect that has clinical relevance.

Absolute with comparative characterization. Mechanism of the immunogenicity, that immunogenicity was discussed, even though it was not the focus of our session, but we discussed it in relation to the quality attributes, that some of the common, the offer from the moderator is that a lot of things were actually learned retrospectively.

Also, the immunogenicity was also raised for the human growth hormone process change, and there was one proposal of process change, even though it was not implemented, but that raised immunogenicity issues.

Comparing the host cell proteins from one manufacturer to another, and host cell proteins more into process, not safety. There was some comment that monitor the host cell protein as a tool to control the manufacturing process and the quality of product, not related to the safety. So that was a comment.

Again, I'm just presenting you the summary of discussions that we had yesterday.

The comment made by the audience and the

moderator, as ground rules within our session, that represented personal opinion.

Acidic forms of monoclonal antibody, succinimide forms, and aggregate literature regarding the effects of biological activity, and there was quite a substantial discussion on the aggregates. Also, glycosylation, leachates, aspartate, the assay has the same charge in the mass as the traditional.

The challenge of comparing quantitative results across independent labs. One example was provided from the audience that from international studies, trying to establish the human growth hormone standard.

So the example provided, from this person, indicated that actually the variability from each laboratory, from different laboratories is about 20 to 30 percent, to identify the mass of their particular standard.

So that raised a concern whether or not, for some of the indication with a narrow clinical index, that 20 percent, 30 percent of difference may make a big difference. So that was one example provided.

Capable of generating ample data, how to

use the data, and you can generate a lot of data, and what those data mean to you, that's another issue.

The example provided is glycosylation, terminal glycans fucosylation that was mentioned from that particular comment.

Physicochemical methods have improved significantly, that we have actually pretty much pretty good consensus on that, and--but still, even though the technology is evolving, but it is still not absolute, and there are limitations remain. Extent is subject to opinion.

Now, there is some consensus that limitations can trigger additional studies. Since that you cannot resolve by the physicochemical characterization, additional tools, additional

means should be used, such as PK, PK/PD, and the clinical evaluation.

Now, some comments from follow-on manufacturer. Might use better methods than the innovator. Not privy to current innovator's methods. Innovator may continue to find new characteristics over time. Some comment on the continued effort to optimize process, manufacturing process.

So it's not something set as stone and is since changing.

Then we move to question three. What are the appropriate standards for the characterization of those identified attributes?

Reed offered the following questions and points of discussion. How to apply comparability concepts without a historical data set; how to link follow-on product laws to the innovator's clinical material without common reference methods of reagents; to what extent does a follow-on protein product manufacturer recharacterize impurities; how to determine that follow-on protein product is

monitoring critical quality attributes.

Counterpoint from Charlie; in most cases, the brand product is appropriate comparator. Acceptance criteria should be based, in part, on brand product variation.

Comments from audience and moderators. Drug product as comparator, and this is a product that goes into people. So there's relatively good consensus that quality attributes related to the drug product that matters, that probably is the material that should be focused on in terms of the quality attributes and the clinical relevance relationship.

Excipients may interfere with analysis of the active pharmaceutical ingredients, and, actually, there are the following two points/counterpoints from the floor.

Extraction may affect API characteristics, which extraction steps are not the manufacturing steps. So it's adding the steps that may affect the active ingredient.

The counterpoint on this is that maybe the

follow-on protein product manufacturer may be able to validate measures by adding or removing the excipients. The one point is that they have made their own active ingredient and then they had the excipients in, so to mimic the situation for a better comparison study.

May need to evaluate intermediates and above drug substance in addition to the drug product. This point was raised in the context of the comparability studies used from the innovator for manufacturer changes.

They raised that the comparability study for the intermediates and above drug substance is a very important part the comparability study and they had a difficult time to see how this part will be evolving for follow-on protein products.

Now, pharmacy, it was pointed out earlier by Charlie, that the final product will be the material for comparison. So these are pharmacy samples, and there are several comments on that.

The number of drug product lots needed for analysis by follow-on protein product manufacturer,

for example, may depend on complexity and purity of product. So how many lots should be used for study was discussed, and there's different opinions on that.

My compare to multiple innovators' product, and follow-on protein product manufacturer does not know how many API batches are represented. Impact on stability, and you receive a product from pharmacy, that product, protein product is relatively unstable as compared to the traditional drug. So that was mentioned.

Reliance solely on limited sample set may lead to specifications that are tighter than innovator's.

Now, there's one comment that actually related to the legal issues and the review issues, so we did not spend a lot of time, but for the purpose of this report, we put it in here just for your information.

Ability of product reviewer to decide on follow-on protein product specification without reference to innovator's proprietary information

and how to deal with that.

So follow-on protein product specifications based on innovator's clinical experience versus analysis of the market products.

This is probably the last slide. There are some additional comments that we captured from the discussion. They are as follows.

Comparability within a manufacturer versus follow-on. In process materials, historical data, clinical experience not available to follow-on protein product manufacturer. That actually was just mentioned in a couple sessions.

Extent of manufacturing changes, incremental versus de novo. Reference standards, monograph available for some products.

Control process to limit modifications, and this was discussed in the context of the heterogeneity of the protein therapeutic product.

Innovator's may stop sharing their experiences if their disclosures are used to support follow-on protein product applications, and that was a concern.

One of the symposiums specifically mentioned is a well characterized biological symposium, which was a symposium that we had very

open forum for discussion.

Industry standards may be harder to identify. Industry thanked for sharing their experiences, where something went wrong. This could help all manufacturers avoid repeating the same mistakes.

That is the end of my presentation. Thank you for your attention.

DR. JONECKIS: In the interest of time and in discussion with my chairpersons, we have a few minutes, if there are any major, and I emphasize major points that people would like to clarify regarding the presentation from the first group. You have a few minutes to go to the microphones and do that.

Again, let me just mention that the transcripts, all comments made are transcribed and will reflect the official discussions that occurred during those various breakouts.

So if there is anything people would like to clarify or make a specific point that was perhaps missed in the presentation from Session A.

Hopefully, then, the group did a fairly good job of summarizing the major points.

Thank you.

The next breakout session is on biological characterization impurities, breakout session B. I drew the shortest straw, and so I will present the summary of the results for today.

Initially, I'd just like to say, on behalf of all the moderators of that session, that is, Janice Brown, Steve Kozlowski of FDA, myself, and Inger Mollerup of Nordisk, and Robin Thorpe of NIBSC, we would like to thank all the participants who actively contributed to the discussion.

As my previous colleague indicated, what we're going to do is capture what we think were the major discussion points and the major conclusions, where there were some, on the various questions that were posed for this session.

The session was started by two

presentations, one by Robin Thorpe and one by Inger Mollerup, the representatives.

Robin made several points, which I won't go into detail here, but just to indicate that the point was that bioassays, depending on what is the need for clinical relevance, that bioassays are often--several bioassays are often developed during development and they serve different purposes, and it's the intention or the purpose of the bioassay or the biological characterization study that one has to consider.

An example was characterization versus lot release, for example. They have different purposes and, therefore, have different intentions.

They may not necessarily, therefore, be suitable for the cross-purposes as to how they are designed and such.

He also pointed out that, in his opinion, clinical efficacy typically requires clinical data. You cannot solely rely upon that bioassay as a result of that, and there were several common findings that he indicated where there was a

failure of a product that had successful results from successful bioassays.

The second point he briefly spoke on was the reference preparations and standards, which ones to use and for what purposes, and, again, it depends upon what is available and the ability to, again, achieve and have an innovator's reference standards for use in the various assays.

He mentioned the ability to use international or pharmacopeia reference standards for bioactivity for the purposes of those assays when they are available.

The last point he said was covered in the morning discussions, plenary sessions, and so no specific comments were made.

This is Inger's slide, and Inger basically posed several questions to help stimulate some thought for the discussion, and those are listed here. What biological characterization, other than safety, would you not find needed for a follow-on biologic and how would you make that decision, and, again, thinking about what is the potency, what are

the various in vivo models of efficacy in PK, what is the mechanism of action or perceived mechanism of action, what can receptor binding tell you and not.

The second, how does formulation affect that; if it is different, how is it different, how is it going to impact the characterization that is useful in the characterization assay.

Lastly, risks; how are we assuming this in terms of all the risks of what is known and not known, in general and specific, for that particular innovator, as well as any particular follow-on product protein; how can you identify these risks and best cope with those risks.

So the first question we addressed, how can clinical relevance of functional biological characterization studies, animals, cellular, binding assays, be established, and as a sub-component or sub-question of that, under what circumstances can biological characterization studies be predictive of efficacy in humans and can this be used to justify limited clinical efficacy

studies.

It was important to note that most people felt that biological characterization studies do provide a measure of biological activity and can complement those physiochemical characterization studies that are also performed on the particular follow-on protein product.

Biological characterization, the overall thought was that biological characterization is not usually predictive of clinical efficacy. Similar to a point that was provided slightly earlier by Robin.

It has to do with the point made that the relevancy of the clinical assay does not necessarily reflect that clinical activity, again, depending upon what the design and intent of that assay is.

In many cases, it can be useful, however, to provide relative differences, can be useful to rule out some differences, and can be more useful to target or better define one's clinical studies.

The second point that was noted is that in

the context of physiological clinical comparisons, in vitro and in vivo biochemical, biocharacterization assays may better define clinical efficacy studies; again, the ability to better select, to better target, to better define what activities--sorry--what types of studies would be necessarily for clinical.

Other major points. If a biological characterization assay or panel of biological characterization assays could be linked to pharmacodynamic parameters, they could be used to justify more limited clinical studies.

Again, if one can have some type of relationship, a clearly established relationship that could be shown to be predictive, this may be useful to, again, reduce the type of extent of clinical studies that would be needed for the particular product.

As a sort of sub-theme of that, if there was an animal model that existed that had a long linkage to clinical efficacy, this model can also be used to justify more limited clinical studies.

This came from some points that were made for some products where there seems to be a long history of clinical linkage to an animal model

relationship.

I should note that this point was debated amongst the moderators last night, but, again, we felt that our intention was to summarize the results of the discussions that were made.

A correlation between the biological characterization assay and the clinical response is not always necessary when performing full or abbreviated clinical efficacy trials.

Again, it gets back to what the biological characterization assay is intended to do in the particular assay and even when it is best to determine and to define what the mechanism of action is or the perceived mechanism of action, its intention for lot release, its intention for characterization are not necessarily the same.

If no clinical studies, efficacy studies, are performed, the biological characterization assay must be linked to clinical relevance. So if

it is possible to, in other words, link and provide evidence that that clinical efficacy study can be performed, it would have to be, clearly, the biological characterization assay would have to be linked to clinical relevance.

Although this point was made, it was felt that it may be rather impractical or highly difficult to achieve this result.

So the conclusions that the moderators felt were drawn is that full biological characterization assays cannot be used in place of clinical efficacy trials. However, in addition to other characterizations, it may be useful to justify limited clinical efficacy studies.

It basically goes to the knowledge that one has about the product, the design of the assay, what it is intended to do and such.

Biological characterization cannot be used to replace safety studies, including immunogenicity, was another major conclusion of that session, that question for that session.

The second question that was addressed,

much like the analytical characterization section, was what are the appropriate standards for the comparison of biological activities.

It was noted that international or pharmacopeia standards may be available for biological characterization assays. Again, that is specifically for biological characterization assays, as our session dealt with that particular topic.

The use of the innovator drug product for comparison of drug substance was clearly a controversial topic, and there were several comments exchanged during both sessions actually on this particular point.

Basically, I think most of the comments resulted around what was feasible and practicable.

Potential concerns on this point for using the drug substance that would be isolated from a drug product was that there could be substantial alterations during purification, and there was a lot of discussion on that; the potential for degradation when one is taking these from

commercially available drug product samples; the ability to accurately measure the various lot-to-lot variability that would be produced in an innovator product.

It was pointed out, for example, that there could be blending of drug substances which would make it harder to determine what the true specification was around any particular quality characteristic or bioactivity characteristic; and, again, trying to identify which drub substance lots were actually used to manufacture the drug product would be very difficult, if not possible to be known by any follow-on protein manufacturer.

It was pointed out, however, that one may be able to control, for example, for some of these types of potential problems.

One example provided was for degradation. So one could potentially control for degradation in that comparator if one used appropriate stability calculations, determined that the rates of degradation were linear, so on and so forth.

It was also pointed out that there was

potential control for drug substance purification reformulation scheme. One of the participants discussed the potential of how to control for isolation of the drug substance in its reformulation. It basically involved the purifying drug substance and reformulating in both the innovator, as well as the developer's or comparator's drug product.

Basically, the net result of this was to establish in-house reference material, and this would provide, in some sense, in their opinion, some added assurance that the material present was truly reflective of the various heterogeneity and variance present in the drug substance.

I'm not clear if this was the same type of example that was discussed during the first session or not, as I think Andy alluded to.

Another point was meted out that there was a substantial amount of potential for copy drift. In other words, I think it was basically stated that the copy of a copy of a copy of a copy is not the original and this would be another concern in

using any type of innovator reference material.

So the conclusions that the moderators felt comfortable drawing from this session was that international or pharmacopeia reference standards should be used for biochemical assay calibration, wherever available, and, again, most of these standards are used, in fact, for the bioassay calibration for innovator products, as well as any follow-on.

Also, that there were clearly difficulties with the use of the innovator drug product in lieu of drug substance as the comparator, and some of those difficulties were previously just described.

So the last question we addressed was based upon biological characteristics, how can product-related impurities be distinguished from product-related substances and from the desired product.

A sub-question from that was if a product-related substance can be identified or distinguished, should acceptance criteria be different for the follow-on product than that

observed for the reference product.

It should be noted that in both sessions we had a lot of comments related to impurities, both product-related and process-related impurities, and the safety thereof.

The moderators chose not to list those because they were not germane to the particular question that was under discussion at this point, but, again, they are clearly reflected in the transcript.

So we drew the conclusions from the comments that were made and pertained to this particular question.

It was considered difficult to distinguish between product-related substances and product-related impurities. Again, I think that was made--the point that was made is that it happens both for the innovator, as well as for the follow-on, and there are various factors that were cited as to why this occurs, even though you are encouraged, under the Q6B specifications document, to, in fact, distinguish and to determine these

various variants, whether they are product-related substance or impurity.

The conclusions were that even if a product-related substance and product-related impurities can be distinguished, it may be of limited value, as not all safety considerations can really be predicted by the activity; that is, the activity of the biological characterization studies.

There were several examples given of where a particular product variant could not have any type of necessarily biological activity and that activity could not be related until clinical studies were performed.

So, again, it was important that some believed that clinical information, in addition to bioassay results, is needed to define a product-related impurity, and this sort of echoes the theme that was also stated, again, in the first question, that, again, you need to make that clear linkage between the bioactivity assay, but, more importantly, that one needs to also determine in

the clinical population what that bioactivity results mean.

Those are pretty much the major points that we concluded from the various sessions, and we are slightly ahead of schedule. So, again, if there are any clarification points on those various conclusions.

Okay. Then I think we will hear from the last breakout on the pharmacology/toxicology group.

DR. EL-HAGE: Good morning. I'm Jeri El-Hage. I'm the supervisory pharmacology in metabolic and endocrine drug products.

Our panel was composed of James Green from Biogen Idec, and Joy Cavagnaro from Access Bio, as our industry panelists; Andrea Weir of CEDR ODE-6 therapeutic proteins, and Mercedes Serabian of CEBR cell and gene therapy, as FDA participants.

I'm going to summarize the discussions of our group. Fortunately, we only were posed with a single question. So I can be brief.

In addition, our session was structured slightly differently since we didn't have a plenary

session yesterday morning. So the agenda was we did a brief background presentation, discussing expectations and considerations for preclinical safety assessments for innovator biologic products, use of relevant animal models, assurance that you're using an animal model in which the biologic is pharmacologically active, considerations of development of neutralizing antibodies which might confound study interpretation, et cetera.

The question that was posed by the organizing committee to our group was in which situations would animal studies be needed and why. Our discussion was organized in a fashion, as suggested by the organizing committee.

We posed examples that had differences in the biologic characterization or biochemical characterization and what animal studies would be needed as a consequence of differences in the standard characterization of the proteins.

In addition, we gave case examples that were based on molecular complexity of the protein products and posed questions around whether

different levels of complexity warranted different extent of preclinical evaluations.

Two of the case examples that we specifically discussed were when the biochemical analyses were not exactly precise, when there were slight differences in purity profiles, degradation product profiles, or when there were slight differences in PK evaluations and what types of preclinical studies would be needed in those specific circumstances.

Fortunately, I think we did reach consensus on many points in discussion, which actually surprised me, based on discussions that have been held internally.

There was a general consensus that preclinical studies are needed for follow-on products; that the in vivo animal studies have increased sensitivity to detect changes. I will acknowledge that there were some minority opinions in the room, that they felt if the biochemical characterization, the biologic characterization showed equivalence, that they felt that studies

weren't needed, but I think the consensus opinion in the room was that they were needed.

One generic manufacturer raised a consideration that we should take into account what we require when an innovator makes process changes as what may be applicable as preclinical studies; that is, if there was a scale-up, change in host cell line, formulation change, what kinds of studies would an innovator do to support safety based on those changes.

There was clear consensus that innovators do conduct preclinical toxicology studies when they make these types of changes.

There was a consensus opinion that the preclinical, whatever preclinical studies may be needed should be designed on a case by case basis, and those studies would be based on knowledge of the innovator product, what the known toxicities of that innovator product is, and what risks are known to be associated with those products, and I will discuss a little bit further, in the next slide, specifics around those cases.

There were also discussion points on why do we need preclinical studies, what do we use them for, and one of the comments was these are used to

write an informed consent form for patients in the study, and many in the room felt that, as has been discussed in the previous overview slides, that a biochemical characterization or a biological potency characterization can't assure in vivo comparability of a product and that a limited preclinical evaluation would provide some reassurance that at least there weren't marked changes in vivo safety profiles for the protein.

Now, obviously, we can't assure clinical safety from a limited preclinical program, but we can at least provide some reassurance that we don't expect marked differences.

There was also consensus that a head-to-head comparison with the innovator product is preferred. It's not absolutely mandatory. There were several discussions around complications of doing a head-to-head comparator study, those being difficulties in obtaining innovator product,

that only clinical formulations of the innovator product are available, that it may be difficult to mimic precisely the formula of the innovator or oftentimes the follow-on product uses a different formulation.

There were other comments and questions around if you need to use a comparator, which one do you use.

The moderators, after the fact, we had some discussion about what we have seen and what has tended to be the case is follow-on manufacturers choose an innovator product that has the most extensive preclinical and clinical safety database. So they have the largest information data set as a frame of reference, but in no way was it implied that that was necessary, but that tends to be what is done.

There was discussion on what types of studies should be done and it was felt that the types of studies should be based on the nature of concern.

I'll backtrack a bit. I think there was a

general consensus in the group, in fact, that there wasn't a difference based on low complexity of the molecule versus high complexity of the molecule.

The consensus opinion was that some preliminary screen was needed regardless of complexity of the molecule. There would also be additional concerns if the biologic had a narrow therapeutic index, if there were known toxicities, and there was also acknowledgment that many biologic products have relatively good safety profiles, minimal toxicity, and we would take into consideration if the known safety profile of the innovator showed minimal toxicity.

There was quite a bit of discussion that, ideally, you could design a single study to look at multiple end points. You could do a single two-week study, four-week study, a bridging toxicology study, per se, and look at PK, PD, local tolerance, and relative immunogenicity.

We understand we will have a discussion today on predictivity of animal immunogenicity for clinical immunogenicity, but we felt at least you

could look--if you did a head-to-head with the comparator, you could look for relative differences in immunogenicity in that study.

There was some discussion that a case example of where this could be done to have minimal impact is for the case example of growth hormone. People do a rat tibial assay as a bioassay and a potency assay for growth hormone, and you could scale up that rat weight gain assay to look at your PD end point, which is weight gain, but also do PK, local tolerance, some limited target organ toxicity based on what is known for the innovator products, and you could use that as a preclinical screening and get more bang for your buck, basically.

Then there was a discussion on what duration of study was appropriate. Many felt that the two-week/four-week bridging study could be the initial screen, especially for compounds with extensive pharmaceutical experience, multiple compounds approved, produced in multiple host cell lines, with extensive clinical experience. Again, the examples could be growth hormone, insulin,

compounds with large therapeutic indices, or low, very good safety profiles, then just a screen with comparators should be adequate.

Then there was a discussion of when longer-term studies might be needed and it was felt that in cases where the toxicity of the innovator was seen both clinically and preclinically, but it took an extended amount of time for that toxicity to be observed; in other words, you only saw the toxicity after three months of treatment in both animals and clinically.

In those cases, longer-term preclinical studies may be warranted to assess comparative toxicity, especially if the clinical toxicity is significant.

To backtrack to our discussions of when you do the comparability assessments and you see differences in PK or differences in toxicity, then there was a general consensus that further studies would be needed to investigate the nature of the differences.

One of the comments was people use

different assays for their PK, is it an assay difference, is it a real difference when you see PK differences, and if there is a difference in biochemical characterization, then you need to characterize why you are seeing that difference and try to assess how clinically meaningful that difference might be.

So the consensus opinion was that there was value added by preclinical safety assessments and the demonstration of comparable safety profile may streamline the clinical program, and, in addition, it provides reassurance of comparable in vivo responses with the follow-on product to what is known about the innovator product.

Again, if there are any follow-up questions or comments, please feel free.

DR. JONECKIS: We are a few minutes ahead of schedule, which is always a good place to be. So what I suggest we do is start right with the immunogenicity plenary sessions.

I would like to introduce Dr. Amy Rosenberg, who is the Director of the Division of

Therapeutic Proteins in the Office of Biotechnology Products.

DR. ROSENBERG: Good morning, all. Because we are in advance of our time, we thought that after both speakers have finished, there would be time for a question-answer session prior to our taking a break.

So good morning and welcome to the immunogenicity plenary session. Immune responses to therapeutic proteins are a problem that impact on both the safety and efficacy of therapeutic protein products. However, considerable controversy exists as to the importance of such responses, as well as to the extent to which they should be investigated, and particularly in the context of follow-on therapeutics.

So in choosing speakers for the immunogenicity session, we sought individuals who not only had a vast experience in both research and regulation of biological therapeutics, but individuals to whom that experience had endowed a very fair and balanced judgment, otherwise known as

wisdom.

So it is a great pleasure for me to introduce our speakers, Dr. Huub Schellekens and Dr. Robin Thorpe. Dr. Thorpe will begin.

Both these gentlemen have CVs that are too extensive to go into any detail as to their accomplishments.

Dr. Thorpe has been head of the Division of Immunology and Endocrinology at the National Institute for Biological Standards and Control in the UK since May 2004, and was previously head of the Division of Immunogenicity at the same institution from 1986.

He has vast interests in cytokines, monoclonals, immunoglobulins, and in the immunology of infectious agents.

His recent interests also include the immunogenicity of biological therapeutics and assay development to assess those.

He is on numerous regulatory committees regarding biological therapeutics and he is an editor for the Journal of Cytokine, and an

editorial board member of the Journal of Immunologic Methods.

So, Dr. Thorpe, it is a pleasure.

DR. THORPE: Thanks very much, Amy, for that very kind introduction. I think you just actually basically said I'm just very old, which is true, which is true. You're too kind to say that.

I would like to thank the organizers for inviting me to give this presentation.

What I have noticed, and I want to say it now, before I forget, is that in your handouts, some of the slides have become scrambled. So if you want the real versions, send me an e-mail or phone me or something and I can send the proper versions to you.

I also refer at some points during my talk to literature references, which, again, are not given in full. If you want those, again, just e-mail me and I can find them.

What I was going to try and do is give a general overview and try and relate what I am going to say to the questions which Amy sent me just

before Christmas; and, thanks, Amy, for that.

DR. ROSENBERG: Merry Christmas.

DR. THORPE: That will result, I think, in my presentation being a little bit sort of erratic, but I can't really do anything about that, but I apologize up front for that.

And towards the end of my talk, I'm going to try and bias towards issues relating to comparability assessment of immunogenicity, which, of course, is the big issue for follow-on products, and I'm going to end up with a couple of perhaps controversial slides, which I would be only too pleased for some of you to disagree with.

So I'm sure we all know what we're talking about when we talk about unwanted immunogenicity, but I thought I would use this slide to point out some points that I think are pertinent.

So what we are actually considering is the scenario where a therapeutic protein is given to a group of responsive patients, presumably does good, and, if you're lucky, nothing else happens and it's fine. But as we are all too aware these days, very

often, in fact, you get induction of antibodies in these patients, which you don't want.

It's not like vaccines, where you want the antibodies. These are unwanted antibodies.

And the consequences of that can be pleiotropic. Adversely, very often, there aren't really any effects or certainly no adverse effects that you can see. So in all cases, certainly, immunogenicity is not important.

Unfortunately, in many cases, it can be, and the kinds of things you can see is antibodies which bind to the therapeutic protein and can alter pharmacokinetics, pharmacodynamics, things like that, which may have implications for responses in the patients to the protein.

Perhaps a worst scenario is where the antibodies don't just bind, but they neutralize biological effects and compromise further therapy, and there are lots and lots of examples of this; Factor 8, interferon alphas, betas, GMCSF, loads of others, monoclonals, and loads of other examples of

this.

The last scenario, which I usually call the horror shop scenario, is where you not only get antibodies which neutralize the products and perhaps some other products, but you get antibodies which potently neutralize all versions of that particular biologic, including endogenously produced substance.

Fortunately, examples of this are relatively rare. The glaring examples, the ones always quoted are EPO and MGDF.

I think perhaps even more important to realize is that the real seriousness of this which causes real concern is where you have a molecule like EPO which has no in-built redundancy. There is no molecule around which can cover for EPO. So if you neutralize its effects, you neutralize the total effects of EPO, and as you all know, if you want red cells, which we all do, you need EPO. So if you wipe EPO out, you don't have red cells, because nothing else can cover for it.

This is not the case for some other

biologicals, for example, GMCSF, which has important biological properties, but if you neutralize GMCSF, it seems that you can overcome those problems, because other cytokines, perhaps things like IL-3, can sort of cover for it.

So this is a less serious scenario and I think this is partly one of the questions that Amy sent me.

So what about immunogenicity? Is it that common? Again, I'm always talking about unwanted immunogenicity here.

If you look in the literature, you certainly find that immunogenicity is not new. It has been around for ages. It has been known for ages, and lots and lots of biologicals can or cannot be immunogenic, and the consequences of the immunogenicity can also be serious or basically nothing. There's a whole scenario between those different options.

I think it's unfortunate, if you actually do look at--this is not complete, by any means, but what you can't really do is make any kind of

prediction about what is causing that immunogenicity and what is going to be immunogenic and what isn't going to be immunogenic and what the consequences are going to be.

So it's a sort of mystery why some of these are immunogenic and why the consequences vary.

However, if you just consider the kind of protein products that are being used, I think you can make some kind of assessment of what you would expect from the immunogenicity perspective.

If you can say that animal-derived proteins, things like calcitonin, things like that, and murine mass in rat monoclonal antibodies, perhaps it is not surprising that these are immunogenic in humans, because they are basically going to be recognized as known cells. So this is exactly perhaps what you would predict.

However, it's now becoming quite clear that it's not just animal sequence proteins that are immunogenic. Human sequence proteins can be immunogenic and potently immunogenic in humans.

There doesn't seem to be a great deal of difference in the immunogenicity, for example, of completely fully human antibodies and humanized

antibodies. So human sequenced proteins are going to be just as much of a problem, from the immunogenicity perspective.

However, with hindsight, I think you can classify proteins concerning their immunogenicity. If you look really hard, and I did, to find these, you can find some biotherapeutics which never seem to have any clear convincing evidence of immunogenicity. The two that I always refer to are gamma interferon and GCSF.

However, Amy pointed out that we already know that to date. I mean, next week, maybe somebody is going to start making antibodies against gamma interferon, because I think the real problem is we don't know why they're not immunogenic. They just don't happen to be. I mean, I've got some ideas, but I've got no proof for them. I don't find any reports of those being immunogenic.

However, it is certainly the case that other proteins can be immunogenic and the induced antibodies can impact on clinical responses or not, and there's loads of examples here, interferon alpha, interferon beta, GMCSF, IL-2, Factor 8, and all sorts of other biologicals.

The rest of this slide just shows the horror shock scenario with EPO and MGDF, and I think limited to those at the moment, where you have very serious clinical outcomes.

So moving on, testing for antibody responses is essential for all sorts of purposes, but it is particularly important for clinical safety of a biological therapeutic, and, also, for product comparability, which is, I think, what we are interested in today.

So how do you do this testing? Well, you need to conduct immunogenicity studies or assessments, and you need to measure antibodies effectively.

I think it is helpful to divide the assays that you use into those that simply measure binding

and those that measure binding and neutralization, because they are quite different.

As I am going to show, you need to adopt panels of assays for assessing the immune responses in the patients being treated with biologicals.

There are a whole range of binding assays which you could measure. You can effectively measure using any immunochemical procedure, but in practice, I think these four listed here are the ones that are mainly used, with some exceptions.

A whole range of different ELISA assay formats have been used to measure binding antibodies, because these are easy to do, usually have high throughput, things like that. So very commonly used to measure binding antibodies.

But, also, other procedures, such as radioimmune precipitation assays are used, particularly with small molecules, but also with things like EPO and surface plasmid resonance, when we're still working with the Biocor machinery, can be used to measure binding, and the realtime kind of characteristic of that procedure, I think, may

actually measure different types of antibodies than those measured in things like ELISAs and other immunochemical procedures.

You can also use procedures like immunoblotting to dissect immune responses to see which components of products are immunogenic or are not immunogenic.

But measuring neutralizing characteristics, you're really stuck with some kind of bioassay, because it's the only thing that is going to measure the biological activity that you will then see neutralized, and you also need to use an appropriate bioassay for the biological question, call it generic assays, and you may need to hone the assay particularly to respond to neutralizing antibodies.

It is certainly true to say each assay has certain advantages and disadvantages, and some of these may be related to the nature of the sample, nature of the antigen, et cetera.

It's not only the assays themselves that confirm these advantages and disadvantages. It's

what you are using them for.

And a compliment of assays is necessary for assessing immunogenicity, because no single assay will give you the full picture and to rely on a single assay might give you an incorrect view of unwanted immunogenicity.

Just to show a couple examples of this, and we were asked to put in examples, so here are some. This basically shows that the binding assays, like ELISAs, don't give you really any indication of neutralizing capacity.

If you look at the top panel, this basically just shows ELISA data for patients that have received GMCSF, and the yellow bars are the ones that neutralize and the pink bars are the ones that only bind.

If you look at, okay, some of the patients that only show binding antibodies have quite low responses, whereas most of the ones, at least in this experiment with neutralizing antibodies, show quite high binding.

But, in fact, some of those patients that

only made binding antibodies had just as high binding antibodies as the ones that had neutralizing capacity. So there is no correlation between ELISA and neutralizing antibody data, and I think that is generally borne out.

If you do the same kind of assessment using Biocor analysis and compare that with the non-neutralizing and neutralizing patient groups, you see what appears to be a better correlation, but I don't think that is actually a true correlation. It's just chance, because if you actually look, again, there is overlap between the non-neutralizing and neutralizing patient groups.

So there is an apparent good correlation between these two different assays, but it's not absolute. So if you want to measure neutralization, you've got to do a neutralizing assay.

So how do you actually assess immunogenicity. As I said, you do immunogenicity studies. It means that you have to develop appropriate assays, and then you have to do

appropriate trials, if you like, or studies in humans.

The only way I can actually show you how you do this is to show an example. So this is an old example that we did years ago looking at GMCSF in carcinoma patients. What we did here was to compare two different GMCSF products, called A and B, and these were very similar. You see I use this term with care, but they were very similar. They were both made in E. coli. They both have human sequence.

So they are not easy to distinguish on the basis of the GMCSF content.

What we did was we did trials in basically similar patients, in the same hospital wards, receiving the same kind of treatment, the same clinical dosing, the same sampling, and then the assays that were carried out, both those carried out to measure antibodies and those to measure clinical correlates were carried out in the same labs, using the same procedures, actually by the same people. So it's completely comparable, which

I think is the important thing to stress.

So if you build in all that comparability, you actually are going to make a valid assessment of the comparability of the immunogenicity.

What you see here, I think you only need to look at the binding and neutralizing antibody results, is an initial indication from the binding studies that both of these materials are pretty immunogenic, and this probably reflects the multiple dosing used and the fact that the patients are not in any way immunosuppressed. They are immunocompetent, so they can make responses.

Product A looks a bit more immunogenic than B. Nineteen out of twenty patients treated with A produced binding antibodies, and about three-quarters of the ones treated with Product B produced binding antibodies.

What is really striking and, I think, unexpected in this study is that it was only patients treated with Product A that made neutralizing antibodies and about 40 percent of these patients did make neutralizing antibodies,

and, as I'm going to show in a minute, in the next slide, these are the important antibodies, from the GMCSF perspective.

No patients receiving the other product ever made neutralizing antibodies, and this was early work. We've subsequently looked at a lot more patients, and none of those have ever made neutralizing responses. So there's something very different going on with these two very similar products from the neutralizing antibody perspective, and that is important clinical because it's the neutralizing antibodies which impacts on the clinical response to GMCSF. They negate it, whereas the non-neutralizing antibodies don't affect it at all.

You can see that on this slide, which shows the ability of GMCSF to mobilize increased leucocyte numbers on the vertical access, and just the cycles of treatment on the bottom axis.

The neutralizing antibodies on the right-hand panel, the non-neutralizing on the left-hand panel. You can see cycle one, when no

antibodies have developed.

Both patient groups can mobilize leucocytes to the same extent in response to GMCSF. But later on, when they've received three or four cycles of treatment, only the non-neutralizing antibody producing patients can still mobilize the leucocytes to the previous level.

The ones with neutralizing antibodies diminished in that respect.

If you look more thoroughly at that study, and, in fact, an expanded version of it, you can find, if you look at the panel of assay results, you can find basically every possible outcome that you might be able to predict.

We were asked to produce real data. This slide shows it, but, basically, the real data is on the left-hand side and you don't really have to look at that, because it is being summarized on the right.

If you just look at patient numbers, and these are just patients that have been selected at random, you can find some, like patients 6 and 14,

which show strong ELISA binding, but weak SPR data and no neutralization at all.

If you look at patient 10, he shows strong binding, strong SBR data, but, again, no neutralization capacity at all. Patient 7 shows strong binding by all techniques and strong neutralization. Patient 2 has strong binding, the binding assays, moderate neutralization.

Patient 11 has strong binding by the ELISA SBR, but weak neutralization. So you can find basically anything you like and you can also see the kinetics reduction, which is shown, it's months of seroconversion, is shown in the brackets on the left-hand table. This varies enormously from very early induction of the, in some cases, quite potential antibodies, whereas all other patients really only respond much later, and, indeed, some techniques seem to be able to pick up antibodies before others.

You can see in patient 8, the SBR data shows seroconversion a month earlier than the ELISA data, which is not reflected in the sensitivity of

the assays.

So it's really basically a case of you will find anything if you look for it hard enough.

This is the kind of data you're going to generate with immunogenicity assessments and you need to be prepared to interpret what often turns out to be a quite complicated scenario. It's not just yes or not, from the immunogenicity perspective. It's usually much more complicated.

I just summarize this on this slide, which is our findings with GMCSF antibodies. Basically, you find everything you can conceive you could find.

You can find antibodies which bind to neutralized GMCSF. They are the important ones from the clinical perspective. You also find antibodies respond and don't neutralize GMCSF. They are not important, from the clinical perspective.

You also find antibodies against non-product-related proteins, host expression system proteins. You find mixtures of antibodies

against products and non-products and if you look hard enough, you even find some patients that never make antibodies, although I think, if you try hard enough, you can induce--you will always induce antibodies against GMCSF because of its immunomodulatory function.

In the questions, it was noted that sub-classes or classes of antibodies might be important from the immunogenicity perspective. So I put this slide in just to address this.

In our experience, and I think if you look in the literature, what you normally find is a pretty classic immune response. You get initial induction of IGM, which may then disappear and switch to a classic IDG response, with IDG-1 predominating.

You might get IGA. I can't really find much evidence of ever really seeing IGE, although there may be some examples I missed from that viewpoint.

I thought I would show this data from

Steve Swanson, Amgen, which shows rather strange subclass distributions in PRCA patients that have been treated with EPO and developed antibodies.

What Steve found, Steve and colleagues, I should say here, is, in some cases, very traditional kind of responses with IDG-1 predominating, as you can see in donor 6 here, but in other patients, there seemed to be a bias towards IDG-4 production, which is rather strange, and this might have some implications for this particular scenario, but I don't think it's general to find this kind of thing. Certainly, we have not found it with GMCSF or interferon, IL-2 or anything like that.

So to move on to how would you actually go about doing the testing, carry out immunogenicity studies, but the common question is how long do I actually have to keep on looking, because if antibodies appear early on, maybe you would not continue to develop that product.

But maybe if you never see any evidence, how much longer do you really have to keep on

looking, and I think, unfortunately, you can't really make any generalities with this because it is product related and it is also dependent on the nature of the disease the patient is suffering with, the schedule of clinical treatment you are using, and so you can't really make any generalities.

However, I think you can say that you're probably going to have to carry out sequential sampling. You're not going to be able to rely on one single sample, because we know that seroconversion time differs enormously.

You also almost certainly are going to be in the problem of needing to carry on looking for antibodies using post-marketing surveillance, because we know that some patients don't make antibodies for years after their treatment has started.

It really is years. Others make very quick responses and, in some cases, those responses may either continue or they may be transient, and the transience may relate to discontinuing therapy,

but there are also instances where therapy is continued, but the antibodies disappear. I'm thinking of enzyme products.

So that's doing classic immunogenicity assessments and as I'm sure you are seeing now, it is quite difficult to do these in humans, and, because of this, there's been a lot of interest in trying to do preclinical work, and I think this is mentioned in the questions.

There's all sorts of approaches you can take for preclinical assessment. You can try and do prediction of immunogenicity. I'm not going to say anything about this, because I don't think there is much evidence that it actually works particularly well.

You can use computer algorithms to identify T and B cell epitopes. This can be very interesting. It can be important.

But what I think you can't use them for is to predict whether or not you'll ever get the antibodies that are going to recognize these epitopes. So not a good predictor of

immunogenicity, per se.

Because of that, people have used all sorts of animal model approaches, because these are obviously easier to use than the human systems.

The animal models are, just as you might expect, going up the kind of species range. Lower mammals have been used, rats and mice, and, also, in some special cases, other animals, like dogs, for certain purposes.

Often, those systems don't work too well, and so people have gone to the obvious extremes of trying non-human primate models and, also, more esoteric systems, which I'm not going to say anything about, because Huub Schellekens is going to talk about it.

So what does the data show with use of animal models? Well, using human proteins in lower mammals isn't really very useful, because, obviously, the lower mammal is going to recognize the human protein as non-self and you're pretty likely to get an immune response, which is not going to be of any use for predicting what is going

to happen in humans.

But these systems can be useful for determining relative immunogenicity, looking at things like formulation, some things like this, and I will show an example of this in a minute.

Some people, to try and get around this problem of recognition of non-self protein, have made animal equivalents of therapeutics or what they consider to be equivalents of therapeutics and they've used these in animals to kind of mimic what might happen in humans, and this, obviously, overcomes the non-self problem.

But what you're looking at is an animal's immune response to an animal protein, and that may or may not mimic what happens in humans, and, again, that could limit this approach.

So the benefit is approach. Predicting immunogenicity in humans is probably limited.

This just shows the example of the kind of relative immunogenicity use, which I mentioned earlier. This is looking at immunogenicity of

human interferon alpha 2A in mice, and this study showed quite nicely that you can use this to look at the effect of route administration and frequency of dosing.

Route administration, I think nothing really unexpected here, shows that IV administration is really basically non-immunogenic, whereas subcutaneous and intraperitoneal administration is much more immunogenic.

The frequency of dosing I think is more interesting, because what this actually shows, if we look at this slide, is it's not the amount of alpha interferon that you give to these mice that is important; it is how often you give the doses. The more frequent the dosing, the more immunogenic was the material.

Another problem with these animal model systems is that the human proteins may not necessarily be active or at least fully active in animal species, and this can be important in some cases, for example, with GMCSF, which is an immunomodulator.

But the thing to remember is that very often, you carry out--you have to carry out toxicity type studies in non-human primates, and it

is quite useful to use these animals for some kind of immunogenicity assessment.

Having said that, you have to remember to design your studies accordingly. Otherwise, you won't be able to interpret the results.

But I think there is a question over animal models for the use of predicting immunogenicity in humans, per se. If you look in the literature, you can find some evidence for useful use of at least primate models. There are good examples with thrombopoietin and things like growth hormone, where monkey models mirrored what happened in humans.

But in other cases, particularly where the products show significant sequence divergence in humans and monkeys, it seems that the prediction is less good and what you tend to see is that the monkey model overestimates the immunogenicity and its importance.

There is one rather strange report using a monoclonal anti-IGE product, which found this to be highly immunogenic in sinos when it was administered by the aerosol route, and this wasn't mirrored when the product was used in humans.

I think, interestingly, it also wasn't

mirrored if you used other routes of delivery into the monkeys, subcutaneous or intravenous.

So I think this is some kind of strange effect which is restricted to this particular product in monkeys, and, again, that may not be very useful from a predictive point of view.

I think this is actually a more striking example of the limitation of monkey models. This is actually what we did with an industrial collaborator, and what we compared was the immunogenicity profile of human GMCSF in monkey models, two different experiments, and humans, and the product is the same and the regime of the administration, the sampling and assays used are, again, all the same.

So you can make a complete comparison

here, a complete valid comparison. What you would see from the monkey model is a remarkable instance of immunogenicity with seven out of eight in both cases producing binding antibodies, and all of those being potent neutralizers.

As you remember from the previous slides, it is the neutralizing antibodies against GMCSF that are important from clinical perspective. So if you did this as an animal model, if you like, for immunogenicity, you would conclude that there are serious problems with immunogenicity.

However, when you look at what actually happened in humans, it was a much milder response. Only about a third of the patients ever made any kind of immune response, and only one patient, that is less than four percent, ever made any neutralizing response.

So the animal model was completely useless in predicting the human immunogenicity and its consequences in humans. So I think you have to be very careful with this approach.

I think the real reason for this is the

immunogenicity is not caused by one factor. It is caused by a whole range of different potential things.

Obviously, the structure of the molecule is important, whether it has novel epitopes, glycosylation can be important, but, certainly, aggregation, degradation, things like oxidation, chemical modification can all be very important, but I think they are the obvious ones.

Perhaps less obvious is product impurities, which can be problematic from their own immunogenicity perspective, but also how they may influence immunogenicity of the product. Also, formulation would be very important, dose, route, frequency of administration, duration of therapy, immune status of patient, the disease they are suffering from, and, often forgotten, immunomodulating properties of the protein itself. All can be important. All can impact and all can be a mixture of factors which are causing the immunogenicity and its consequences.

So I think, to finish off now, I hope I

have convinced you that if you have a biological, whether a follow-on or a new molecule, if you like, you would need to look into immunogenicity and its consequences.

If you don't want to take notice of me, which is entirely reasonable, I think you're still going to have to do those, because the regulatory agency published views all suggests that you should do this, and there's a whole range of documents that you need to look at if you're going to be involved in this. This is just a couple on this slide.

Moving on, to finish off with actual comparability, and this would be comparative immunogenicity, I think there are important points that you really have to take into account if you're going to be involved in doing this, and, of course, this is what is going to be necessary for follow-on products.

A real no-no is to look at published data for particular products and assume that your product is going to show the same immunogenicity.

It might, but it might not, and the only way you're going to address this lack of certainty is by doing a study.

Equally valid is the conclusions on immunogenicity of products obtained by comparing data from different studies, using different products and perhaps different assays, are also usually invalid, because there are other factors in these kinds of studies that may or may not impact on immunogenicity rather than the actual comparison of the two products.

Full comparative immunogenicity, you need to remember that what you are actually doing and what you need to do is design studies to demonstrate whether or not the immunogenicity of two or more products is the same or significantly different.

This may affect and, in fact, nearly always will affect the design of the studies and particularly their interpretation. You also need to remember that you don't just need to measure the antibodies and measure their characteristics.

You need to look into the consequences of the immunogenicity in patients with appropriate procedures.

The fact that there is available immunogenicity data of a marketed product doesn't influence the need for comparative immunogenicity studies, because as I showed earlier on, you can take what appear to be two apparently identical products and they show quite different immunogenicity, and you could only assess that by comparing the immunogenicity. You can't make any assumptions from the marketed product.

However, I think it is important to be prepared for what you're going to do with the data that you generate with this comparative immunogenicity.

If you're lucky and what you show is that the immunogenicity profile and its consequences is the same for the follow-on as it is for the initiator, you're fine. That's great. You've got the ideal scenario.

If you find that the immunogenicity of the

follow-on is dramatically greater than the innovator product, the consequences are important, I think everybody would agree that there are serious problems there.

What about if you actually find that the immunogenicity profile is lower in the follow-on than the innovator product? Well, it's fine from the immunogenicity point of view, because you don't have immunogenicity problems; maybe even claim that the follow-on is better than the innovator.

But I think the problem is you would have to explain why there is a difference and it would be difficult not to do that on some basis that the two materials are dissimilar.

So I think you are going to have to be very careful if you ever find yourself in that scenario. I actually have met somebody who did find themselves in that scenario, and I wasn't convinced of his explanations.

I will finish off by just acknowledging the people in my lab, particularly Meenu Wadhwa, Chris Bird and Paula Dilger, who do the assays;

Rose Gaines Das, who does the statistical analysis; and a whole range of collaborators who we have worked with this on this topic over the past 12 years or so.

Thanks very much.

[Applause.]

DR. ROSENBERG: Thank you very much, Dr. Thorpe. That was a very informative talk.

Our next speaker is Dr. Huub Schellekens. He is the Director of the Central Laboratory Animal Institute of Utrecht University, and he is on the faculty of Pharmaceutical Sciences.

He spent 30 years in preclinical and early clinical development of therapeutic protein products and he has become recently interested in the immunogenicity of therapeutic proteins as a result of his interest in therapeutic protein products.

He is now concentrating on the predictive animal models and assay development.

So welcome, Huub.

DR. SCHELLEKENS: Thank you, Amy. Good

morning, ladies and gentlemen. Thank you for the invitation to speak about what started as a hobby, the question why patients make antibodies to therapeutic proteins, that by now has become nearly a full-time profession.

I will have a little bit other perspective than Robin. I will try to approach the problem more from the clinical point of view and, also, the studies we are doing at the Utrecht University, where we are interested in the question why patients make antibodies to therapeutic proteins.

As you know, immunogenicity is considered the key issue for bio similars. First, let me acknowledge my colleagues at Utrecht University, two people from the pharmaceutical faculty, Daan Crommelin and Wim Jiskoot, who have specialized in proteins. I will show you how they helped this program. Also, Suzanne Hermeling, who is actually doing the animal experiments.

Immunogenicity and biotech comparability, I think, is the heart of the bio similar discussion, because it shows that we are not able

to predict with the current analytical methods what the biological properties of these proteins will be, and we know that the immune system, as I will show you, can detect changes or differences in proteins that somebody in the lab cannot see.

Moreover, probably the most important issue, that immunogenicity is, in some cases, not a trivial issue. It can have serious clinical consequences.

Well, first, I think if we look at the literature on induction of antibodies by therapeutic proteins, I think that we can say that nearly all therapeutic proteins will induce antibodies. I only know one example, which is GCSF, to which nobody has ever reported antibodies.

I think there are reports on antibodies to gamma interferon.

The incidence, however, differs. It can be very rare, like in the case of EPO, but it can also be very frequent. I think the majority of patients who are treated with beta interferon derive from E. coli produced antibodies, and I

think it's very, very important to realizes, if you think about causes and think about ways to solve the problem of immunogenicity, that there are two different mechanisms by which patients make antibodies.

There is the classical reaction to new antigens, but in the majority of cases, what we are seeing in patients who are being treated with therapeutic proteins, there is a different immunological mechanism; that is, breakdown of immune tolerance, breakdown of B cell tolerance, and that is a mechanism that is not completely understood yet.

There are major differences between the two types of immune reactions we can see in patients. If we look to the classical reaction to foreign proteins, I think the prime example here would be streptokinase, a product from microbial origin.

If you give this to a patient, he will produce antibodies very fast. Often, a single injection is enough and it is mostly neutralizing

antibodies that do away with the efficacy of the product, and, of course, it is easily explainable. It's the foreign antigens.

But in the case of breaking B cell tolerance, that is another mechanism. That is a very slow mechanism. In general, if you look at the different products, it takes patients at least six months to develop these antibodies.

Some do it after a year or two years. I think the cases in which patients only produce antibodies after a few years are pretty rare.

So you need long treatment. In general, these antibodies are binding antibodies and these antibodies disappear after treatment, and I think there is some convincing evidence, also, that in certain cases, the antibodies disappear upon prolonged treatment.

We know the cause, more or less. I will go into more detail on the causes of antibodies, but in the majority of cases, it is impurities and aggregates, and that also points to the two different immunological mechanisms we think that

play a role and the impurities may act as a second signal, as a danger signal in activating the mechanism that is breaking B cell tolerance.

In the case of aggregates, we think it's the physical way the protein is exposed to the immune system; that there is evidence that the B cell receptor is also sensitive to the three-dimensional structure or the repeated structures of proteins that can be present in aggregates.

These are the factors, most of the factors that influence immunogenicity, and, as you will already see, I will go into more detail on these factors, the structural properties are the minority of the factors.

The main factors are somewhere else and are non-structural.

Let's go to the structural properties. The first structural property, the degree of non-self I already discussed. That plays a role in the antibody induction by things like streptokinase.

Let's go to glycosylation. There is no example yet, there may be examples in the future, but there is no example yet in which a change in

glycosylation or hyperglycosylation has led to a problem of immunogenicity.

The only convincing evidence we have on the influence of glycosylation on immunogenicity is the lack of glycosylation and the examples are GMCSF and beta interferon, and, in the case of beta interferon, I think there is pretty convincing evidence that has to do with the solubility.

If you have a non-glycosylated beta interferon, it is more hydrophobic, it is less soluble, you get aggregates and you get antibodies.

I will come back to the non-glycosylated beta interferon later in my discussion on animal models.

Other factors. An important issue that Robin didn't mention in the assays, these assays are not standardized. So that makes it impossible to compare results from one lab to the other.

Some years ago, we sent around a blind

panel of antibodies to alpha interferon to the different European labs specializing in antibody testing and when we got back the results, there was more than a 200-fold difference in reported activity. That makes it impossible to compare the results of your trial with published literature.

There is some activity now in Europe, in the regulatory environment, to get assays standardized and we now, more or less, have a common assay for antibodies to beta interferon and there is also some activity in the area of antibodies to EPO.

But, please, realize that an important issue will be also with the bio similars and with the regulatory aspects of bio similars, the standardization of the assays for antibodies.

Other factors. Formulation, this is a classical case. What you see here is different formulations of interferon alpha 2A. Let's concentrate on the highly immunogenic formulation A. This is a freeze-dried HSA contained formulation that, according to the manufacturer,

you can keep at room temperature.

If you kept this at room temperature, something happens. There are different screens, but concentrate on MO. MO is an oxidized form of alpha interferon that appears in the formulation if you keep this at room temperature.

This is a reactive molecule. It reacts with the alpha interferon and also reacts with the HSA in the preparation.

So what you got here was complex of HSA and alpha interferon which was the cross of the high immunogenicity. The formulation was changed. HSA was taken out. It was a liquid formulation that has to be kept in the fridge, and the extra problem of immunogenicity has disappeared from this preparation.

Other factors. Downstream processing. An example, with Factor 8, a new pasteurization step introduced, and problems with immunogenicity. Impurities and contaminants, already mentioned.

If you look at the literature of insulin and growth hormone, you see, more or less, the

immunogenicity which was reported in the early years disappear and there is a high degree of relation with purification with these products.

Duration of treatment I already mentioned. Duration of treatment is a factor in itself, because you could argue, if you treat the patient longer, you give him more protein and if you give him more protein, he is more likely to produce antibodies.

But if you look at the interferon betas, which is different considerably in specific activity, they more or less produce the antibodies in the same period of time, independent of the amount of protein they got.

So duration of treatment is a factor in itself.

Other factors. Route of administration. Put all literature together, make a rank order, this would be the rank order of the most immunogenic route to the least immunogenic route.

That means if you are going to do trials, you have to compare the right route of

administration. That is also true for the type of disease. There is convincing evidence, for instance, in alpha interferon, and we know now with EPO and Eprax in patients, that patients with cancer are less likely to produce antibodies than patients with other types of diseases.

So you cannot extrapolate your toxicity or immunogenicity data obtained from one group of patients with another group of patients.

Genetic background of patients will not help you. There have been large studies in insulin, for instance. I know that Nicole Casa du Fal looked at aplasia patients. There was no relation with immunogenetic background of the patient.

The only genetic background we are sure of is in hemophilia, because the Factor 8 is a defect, a type of defect in the gene that more or less predicts whether you will make antibodies or not.

So if you are doing trials with Factor 8, you have to normalize your group of patients for the genetic backgrounds of the disease, and, of

course, that's not nice for you, but it's nice for us in the lab, because it keeps the grant money coming in.

There are still unknown factors in immunogenicity and the prime example has been shown many times before.

This is Avonex. Avonex produced at different production sites. What you see on the left is the production site in Germany. What you see on the right is the production site in the United States. You see here large differences in immunogenicity and up til now, this difference has not been explained.

I know that Biogen has looked in more or less everything they could look at, aggregates, oxidation, whatever, and up til now, they've never come up with a satisfactory explanation why these two different preparations differ in immunogenicity.

Let's go to the consequences of antibodies. Robin already showed you this slide, more or less, what happens. If you take every

antibody results and put them together, I think the general conclusion is that in the majority of cases, there are no clinical consequences of binding antibodies.

But if you look at neutralizing antibodies, then the case becomes different. I don't know of any example of patients making sufficient amount of neutralizing antibodies without any clinical effect, and the most likely clinical effect is the loss of efficacy and, of course, the most feared clinical effect is the neutralization of an endogenous protein and if that endogenous protein has an important biological function, the patient gets into trouble. I will show you examples with the MGDF and with the EPO.

I have to warn you, though, Robin mentioned the interferon. It seems that the interferon antibodies have no other effect than reducing the clinical efficacy, but I predict that in the end, we will see clinical consequences of the neutralization of the endogenous beta interferon in patients with high antibody levels to

beta interferon.

There is only one single beta interferon gene that has been preserved through evolution. It is nearly true for all species that they only have one beta interferon gene.

There must be a very specific reason why we have one beta interferon gene, and I cannot imagine that you can neutralize the activity of this gene without any harmful effects in the long run.

But let's go to the loss of efficacy. This is an example of hepatitis C antibodies to interferon alpha, and you see here high levels of antibodies do away with the sustained response.

Let's go to the neutralization of the native protein, already mentioned by Robin. This is the example of the mega growth and development factor, a pegulated molecule, by the way, in development by Amgen. They stopped development because they saw, in clinical trials and in a volunteer study, the patient produced antibodies to the product, cross-reacted with endogenous DPO.

You need DPO to make block platelets. So without this factor, the patients became severely thrombocytopenic and needed platelet through infusions and some for a prolonged period of time.

Let's go to the epoetin and alpha and pure red cell aplasia cases. You are all aware of the problem that started in Europe and Australia and Canada in 1998 with the introduction of a new formulation of Eprax.

In fact, the European authorities, because of the scare of BSE and HIV, requested manufacturers to take out all human proteins from their formulation, and Johnson & Johnson obliged and they changed the formulation from human serum albumin to polysorbate 80 and glycine, and since that period, we have seen an upsurge of cases of pure red cell aplasia.

There are a number of explanations going around on what, in fact, has happened. We have, in Utrecht, made the hypothesis that we have found EPO associated with mice cells in the preparations and they could form the molecular structure that the

immune system would recognize and would break B cell tolerance.

There are other explanations also offered, like direct interaction between tween and EPO. Silicon, I don't think anybody really believes that silicon has an effect on immunogenicity, and Johnson & Johnson has come with data on leachates from rubber stoppers which may play a role. Also, mishandling, which I think in individual cases may explain what has happened with EPO and pure red cell aplasia.

But let's go to prediction, which is probably the thing you are most interested in. This is, more or less, what you can do to predict immunogenicity. Sometimes you can on the facts, but whether the product is pure or not, to make some predictions on immunogenicity.

Epitope analysis. There are a number of companies that are suggesting send me your sequence, and I will tell you whether your product is immunogenic or not.

I think this will work in products from

non-human origin, but as I showed you, in the case of breaking B cell tolerance, that has nothing to do with the sequence. That is in the formulation, that is in the aggregates, and that is not predictable by epitope analysis, whether this will happen with patients or not.

You also see sometimes reaction with patient sera. So what manufacturers do, they take sera of patients who have produced antibodies and look at the reaction to their product and then make conclusions. That's a fallacy, because that you are testing there is antigenicity and antigenicity is only the property of a molecule to bind an antibody.

That doesn't predict whether this product will include the antibodies or not in patients.

I think there is beautiful literature on measles vaccines, looking at the epitopes that react with neutralizing sera and then using these epitopes to induce antibodies, and then there is hardly any relation between the two.

But let's go to the animal experiments,

already discussed a little bit, conventional animals. I think they can be used, more or less, for relative immunogenicity, but only if there are gross differences between preparations.

I think for subtle differences between preparations, there is no way that a conventional animal will show this difference.

Non-human primates. I spent 15 years of my life in the primate center. I probably still have the world's record in sticking these products in rhesus monkeys and in chimpanzees, and, intuitively, we think they're more like humans, but I have examples of products that were immunogenic in primates which were not immunogenic in patients, and the other way around.

We still believe that a way to go could be the immune tolerant transgenic mice, and I will show you examples, but let's go first to the prediction on the level of purity.

One of the hobbies we have is the collection of these products for all parts of the world and what you see here is EPOs collected from

India, Korea, South America, and comparing their--this is--I forgot the type of analysis here, but it shows the differences in the bands and what is most frightening here is that you see differences between different lots of the same manufacturer.

Of course, you cannot predict absolutely on the basis of this whether there will be safety issues here, but you can suspect them.

Let's go to the transgenic mice. We recently made a mouse which is immune tolerant for human interferon beta. So that's more or less like a human patient, from an immunological point of view.

I will save you the construction of the animal, but we did all kinds of controls to show that they had the right construct and whatever. It's behind murine interferon beta promoter.

The problem we have in the lab is that they're not very fertile. So we have a hard time breeding them, but that's a solvable problem.

Interferon beta was already discussed

yesterday, the difference between the two CHO-derived products, but what I will show you is the difference between the CHO-derived product and the E. coli-derived product.

I already mentioned the lack of glycosylation makes it less soluble and gives aggregates, and this is the highly immunogenic preparation in patients and that we can reproduce in these mice.

What you see here is the CHO interferon beta in wild type mice. As you would expect, these animals produce antibodies.

If you give the same preparation to the transgenics, you see here that they are immune tolerant to the CHO-derived protein.

If you give them the E. coli-derived protein, the same amount of protein, you see that the tolerance in these animals is broken and that they produce antibodies. I think this is also evidence that what we see in patients, if we give them E. coli-derived beta interferon, is, from an immunological point of view, different than what we

see in the limited number of patients who make antibodies after they are treated with Avonex.

What we also did is we revived literally the mice which were used by Roche to study their problem they had with this highly immunogenic freeze-dried HSA-containing preparation. They were in the deep freezer somewhere in Basil, and so we had to revive them.

What we are doing now, because all these models up til now have been used in a yes or no fashion, so in a qualitative fashion, do aggregates break immune tolerance, but what we are trying to do is whether we can really make this into a quantitative model.

But I will introduce a caveat here, which also, in my view, will have implications for what we do in human patients. But let's go on to the type of experiments we do.

This is the wild type mice who have been treated with different alpha interferons and, also, alpha interferons which were mistreated, mistreated to get them oxidated or mistreated to get them into

aggregation.

What you see here is their immunogenicity in wild type mice.

If you look at the immunogenicity immune tolerant mice, you will see that only the metal catalyzed product seems to be able to break tolerance and induce antibodies in these mice, but there is a caveat here and the caveat is what you are testing for.

Let's go to this slide. This is more or less the repeat of what I have already showed you, but what we are testing here is antibodies to the genuine alpha interferon 2B. So an unaltered interferon alpha 2B.

Let's concentrate on G, which is the glutaraldehyde cross-linked material, and on B, which is boiled alpha interferon. We thought boiling was a way to precipitate the material.

This is what you see in the wild type mice. But, now, let's go to the results if you do not test for antibodies to interferon alpha 2B, but, in fact, to the preparation you have put into

the mice.

So the mistreated alpha interferon, standard looks different. You see that boiling does away completely with the alpha interferon. There is nothing there to react and there is nothing there to induce any antibodies.

But if you look at, again, the glutaraldehyde, you will see that we didn't see any activity with the alpha interferon itself, but there is activity against the modified interferon.

So the lesson here, I think the lesson, also, for if you do the clinical trials, that if you have an impurity or a degradation or whatever in the product, not only test, and if you are interested in immunogenicity, don't only test for the actual product, but also test for the antibodies against the modified material.

My conclusion. My conclusion is that immunogenicity can have serious clinical consequences. It is not yet possible to fully predict the induction of antibodies, not even with transgenic mice. We need to do much more on the

quantitative validation of these models to have them ready to test biosimilars, and that leaves that the only way for relative immunogenicity, if you have to compare different preparations, is to do clinical studies, either pre- or post-registration.

Thank you very much.

[Applause.]

DR. ROSENBERG: I think there is time now for some questions and responses. Dr. Thorpe, would you please come up and join us?

There are microphones between most of the aisles, and we welcome specific questions. Please direct them to the particular speaker you would like to answer.

DR. SENSABAUGH: Good morning. I'm Suzanne Sensabaugh, from Sicor, Inc., a subsidiary of TEVA Pharmaceuticals.

I would like to thank both Robin and Huub for giving us very wonderful talks this morning. You provided us with a lot of data and information on immunogenicity. I would like to thank Amy for

bringing you here to us today to give these talks.

The data and information that you gave us was a wealth of information. We heard that GCSF has not been shown to be immunogenic. We have heard where GMCSF has shown to be immunogenic.

We have also discussed some product factors relating to immunogenicity.

The FDA issued a guidance document on assessing--taking a risk-based approach to immunogenicity. What I would like to ask you to do is, if you could, take some of that information and data that you have provided to us and sort of address this approach.

Amy, maybe you would be the better person to do that, considering that you're within the FDA.

DR. ROSENBERG: Can you just tell me, which document were you referring to?

DR. SENSABAUGH: This is a guidance document. It's on immunogenicity and taking a risk-based approach. I don't remember the name, off the top of my head.

DR. ROSENBERG: A guidance document?

DR. SENSABAUGH: Yes.

DR. ROSENBERG: Is anyone else from FDA here aware of the issuance of a guidance document

coming from the agency? I mean, several articles have been written.

DR. SENSABAUGH: Right. Okay. I apologizes. I'm sorry.

DR. ROSENBERG: But there has been no guidance document.

DR. SENSABAUGH: Okay. Maybe we can talk afterwards.

DR. ROSENBERG: Okay.

DR. SENSABAUGH: Okay. Sorry.

DR. ROSENBERG: Thanks.

FROM THE AUDIENCE: I have a question regarding Dr. Schellekens' remarks, but would welcome an answer from anybody.

You spoke about a lot of interest, and I know there has been particularly in Europe, in creating standardized assays so that you could compare across products, potentially, and across studies.

But in your last slides, you also showed that you can get antibodies that are specific for a particular variant form or degradation product of a protein.

So given that, it would seem that there is a significant limitation to a common--if you

compared two products, one may make more antibodies to one variant and the other may make more antibodies to another variant, and if you have a common assay, they may look the same or different, but it won't tell you really the story at all.

Don't you need to test each product for its ability to induce antibodies to itself and to the various forms it may take in the body or after storage?

DR. SCHELLEKENS: I think this will be an issue of clinical relevance. If the antibodies to the impurity have no biological or clinical consequence, then I don't think there is any need for standardization.

I think in the case of beta interferon, what drives the beta interferon is a large number

of patients are being treated with beta interferon, but have such a high level of neutralizing antibodies, that it is a waste of the drug.

There I think the effort is, first, the effort to get some kind of normalization in the statements on the immunogenicity of the preparations in the package insert, but maybe, in the end, to be able to put a clinical relevance titer into a package insert or into an advice to clinicians who are treating patients with beta interferon.

I think it depends on the clinical relevance. If there is--

FROM THE AUDIENCE: Well, could you envision a situation, say, with alpha interferons, where product A induces a lot of antibody against product A and product B induces a lot of antibody against product B, but they don't necessarily induce as much of that cross-reaction?

Those would be relevant antibodies, because they would neutralize or change the PD of the product, but not necessarily--

DR. SCHELLEKENS: In that case--

FROM THE AUDIENCE: So then a common assay is going to limit your--

DR. THORPE: That definitely happens. We've got GMCSF and other things and that does happen. But in the end--

FROM THE AUDIENCE: Right. So you have to be cautious about common assays, because you could bias it to favor one or the other, but it wouldn't tell you the story.

DR. THORPE: That's right. In the study that Huub started to mention, the beta interferon, it's taken so long, because you have to do so many combinations of antibodies with different antigens and what you see is sometimes the antibodies cross-react completely, sometimes it's partial, sometimes they don't.

So you see everything again, and that's why it's taken four years, is it now?

DR. TANIGUCHI: Gary Taniguchi, BioMarin Pharmaceutical. A question about--you gave two mechanisms, but what about--have you observed any

cases where people had preexisting antibodies, like preexisting antibodies to FAB or FAB Prime 2s, for instance, or even possibly EPO, that existed prior to the patient being given the molecule?

Have you seen any clinical--I guess, an increase in the antibodies from those cases where the memory B cells react?

DR. ROSENBERG: Could you--people could not hear you. Could you speak louder?

DR. TANIGUCHI: Gary Taniguchi, of BioMarin. The question is about preexisting antibodies. Have you seen any cases, like, for instance, in the EPO cases, where prior to patients given therapy, and, also, there is preexisting FAB and FAB Prime 2 antibodies, have you seen any cases where that has caused memory B cell activation and things like that?

DR. SCHELLEKENS: I think there is only one single example of a patient with natural antibodies to EPO, which was also discovered by Nicole Casa du Fal. In all other cases, I think there is pretty much evidence there were no

preexisting antibodies to EPO.

I know that in other cases, there are people claiming that many individuals already have antibodies to things like alpha interferon and IL-1, but that's highly disputed.

I think that Roche, who has the largest collection of sera and patients, they screen for antibody for alpha interferon, never found a single patient with preexisting antibodies.

DR. THORPE: I would agree with Huub. There are a lot of reports of what are described as preexisting antibodies which are probably not real antibodies that have been induced by anything that relates to the antigen. They're either artifactual or they have been induced against other things, perhaps infectious agents, and they cross-react.

But I think there are some strange--well, anyway, against GMCSF, do seem to occur not in everybody, certainly, but in a significant number of people, and I think it's case by case. You can't make any assumptions about that.

I'm not aware of patients that have had

preexisting antibodies, for example, against EPO and things like that which have been boosted when they've received them.

It may happen, but I don't know whether it does.

DR. ROSENBERG: Terry?

DR. GERRARD: Terry Gerrard. I have two comments. The first is you have talked about a lot of the factors that can influence immunogenicity, but is there a common factor that goes through some of these things? Like you talked about glycosylation of the beta interferons, but it really boiled down to solubility and aggregation.

The same with the alpha interferons. The problem was leachability or something in formulation caused aggregation.

So how much of this, when we talk about different factors, whether it's glycosylation, removal of HSA, leachates, whatever, how much of it distills down to solubility aggregation, a factor which certainly can be monitored?

DR. THORPE: I think aggregation is

certainly important and if it's in a self-aggregation, aggregation of the product with other things, perhaps excipients, certainly, aggregates show enhanced immunogenicity compared to some non-aggregated material.

But I think there are other factors that may be important. Certainly, the patients that you are treating, their underlying disease, whether they be immunosuppressed, either purposely or not, and all the other things on those lists I think are important.

A lot of them could relate to aggregation, such as the formulation problems that I had up. I think there are some examples of that where it definitely is aggregation of active substance with incipients.

I think aggregation is important, but there are other things.

DR. ROSENBERG: I think one thing that hasn't come out is the fact that we are not inherently tolerant to all our self-proteins, and particularly self-proteins that are present at very

low levels in the circulation.

So I don't think there was any evidence for the peg MGDF that it was an aggregate problem. Basically, that's a factor that is present at very, very low concentration, and, in fact, preexisting antibodies were seen to that, although, again, I don't know how specific those were, because I didn't see the competition assays.

But it would not be surprising if those were, in fact, real preexisting antibodies. So that is another factor that certainly comes into play and the adjuvant activity of impurities that might trigger antigen presenting activity through toll-like receptors, I think, is also something that is very important.

MR. CORIN: Gene Corin, Amgen. I have two questions for Dr. Schellekens. Excellent, talk, by the way, but there are always questions.

First, if I'm not mistaken, you stated that the genetic background of patients doesn't seem to make any difference in immune response or to be a factor.

Do you really think that there is enough data and literature to substantiate this? I'm not aware of any extensive HLA typing, for example, in

patients who did develop antibodies versus those who didn't. So that's one.

The second, in your immunotolerant mouse model, these animals are tolerant towards the products that you make them transgenic for, but their immune system is still a mouse immune system and I would say that these models would be still only good for these gross differences in antigens.

DR. SCHELLEKENS: Regarding the immunogenetic background of patients, the only statement I wanted to make is we have no evidence about the relation between the immunogenetic background of patients and the fact that they make antibodies.

I refer to published literature, I think, on insulin a number of years ago and there are non-published data on--I know Roche looked at their alpha interferon patients and the data from Nicole Casa du Fal. I don't think she published them

either, and she doesn't seem to find a relation between their immunogenetic background and the production of antibodies.

I completely agree with you on the mouse models and I tried to be a little bit cautious that we are still working on these models and, from a quantitative perspective, we don't know how predictive they are and they will still be mice. They will not be human patients.

DR. ROSENBERG: With respect to that, Huub, the transgenic mice, even though you put them under an endogenous promoter, they may have different copy numbers. They may cause the protein to be expressed in a way that is different from the way the protein is normally expressed, plus interferon beta, if I'm not mistaken, is species-specific, so it's not going to have its immunomodulatory activity in the mouse.

I mean, I think those are a few of the subtle points of difference between the transgenic model and as it reflects on their capacity to predict.

Andrew?

DR. CHANG: Andrew Chang, from FDA. I just want to offer one comment regarding one of the

questions that came up earlier regarding the preexisting antibody in normal people.

We had one situation, and there is some literature reported on preexisting antibody against thrombin, which is a final protease, in a cascade, and those patients, after receiving the topical thrombin, that they generated a high titer of the antibody for that.

So I have a question to Dr. Schellekens. I hope I pronounced your name right.

Now, you made a comment regarding using animal models to predict the differences between the two different product preparations, for example, not use an animal model to predict the immunogenicity to predict whether there are any differences.

You mentioned that you need drastic differences in order to see that. I'm just wondering whether there is some better animal model

that is very sensitive. The example I can come up with is the animal model to Factor 8.

An animal, like a CA57 Block 6 normal mouse is very insensitive to Factor 8 in terms of immunogenicity, but after you knock out Factor 8 gene from those mice, they become to be very sensitive.

Any comment from you on that?

DR. SCHELLEKENS: It's an interesting model.

MS. GRAHAMS: I'm Imogene Grahams, from Regenerant. That was an excellent session, by the way.

In a clinical study, if you have randomized placebo control design and within the active arm, some of the patients develop antibodies and some don't, it is tempting to try to compare the group who develop antibodies with the group who don't, but that is not a proper comparison, because it is fraught with limitations of subgroup analysis.

My question is do you have a

recommendation for a proper clinical trial design to evaluate the effects of the development of antibodies?

DR. THORPE: I think it is very difficult to make generalities. It depends on certainly the product you are going to use and the patient group.

There are papers that have been published to try and address this, but I think the overall conclusion that most of them come to is that it is all case by case and there are some generalities which are glaringly obvious, like you have to sample appropriately, things like that, but there are real imponderables left to case by case.

I can't actually make any comment more than that. I'm sorry.

DR. ROSENBERG: Huub?

DR. SCHELLEKENS: I agree.

DR. BEN-MAIMON: I'm Carol Ben-Maimon, from Barr Pharmaceuticals. I have a point of clarification from Dr. Schellekens' talk and then just one question.

You presented a slide on Avonex and you

pointed out that there were differences in the antibody formation and those between the two products, one manufactured in Europe and one, I think you said, here in the United States.

My question is, are both of those products on the market?

DR. SCHELLEKENS: No. I think the American product is on the market.

DR. BEN-MAIMON: I'm sorry. Which?

DR. SCHELLEKENS: The American.

DR. BEN-MAIMON: The American product is on the market.

DR. SCHELLEKENS: I think the stuff that was made in Germany was only used in clinical trials. Is that correct?

DR. BEN-MAIMON: So the manufactured product in Europe that was the basis for the approval from the clinical trials was not the product that is currently manufactured in the United States.

DR. SCHELLEKENS: No. It--

DR. ROSENBERG: That is correct.

DR. BEN-MAIMON: So I would just raise the question of the clinical relevance of those findings. But that really wasn't my question.

The other question I think I wanted to ask was could you comment on--and I think in Dr. Thorpe's presentation and in yours, there were some presentations on patients who formed antibodies, both animal, but there were also some human studies, and there were some very limited numbers.

Could you comment on the numbers of patients that can--clearly, in some instances, you can see differences with very limited numbers of patients, and comment on some of the impact of that.

DR. SCHELLEKENS: I think that's clear. I mean, if you want to study beta interferons, for instance, if you want to make a similar derived beta interferon, I don't think you need so much patients to show whether your product is as immunogenic as the innovator product.

I think in the case of EPO, it will be very, very difficult to do an immunogenicity study

and doing clinical development. I think in that case, you have to rely on post-market conveyance to ensure that the product is safe.

So I think it depends on the relative immunogenicity of the product you see in patients.

DR. ROSENBERG: I just wasn't sure about what your comment was about the clinical relevance of the antibodies. This was a situation in which those products were deemed comparable by bioanalytical techniques and then when tested for immunogenicity were found to be different.

FROM THE AUDIENCE: I think I was alluding to the fact that through comparability protocols, that product was approved here, and there clearly were differences in immunogenicity in the clinical trials, but it wasn't deemed necessary to do any further clinical testing.

I think the same sort of comment can be made to the Eprax situation, where you do need large numbers of patients to see what are probably subtle differences and small clinical impacts, but the fact of the matter is brand companies are using

comparability protocols to make changes that, in that case, clearly resulted in a clinical effect, but are not going ahead and doing tens of thousands of patients.

DR. ROSENBERG: But in this case, the immunogenicity was, luckily, less, and that's why follow-up studies were not deemed necessary.

I think if the study had shown more, we would have had to consider whether additional safety and efficacy studies were necessary, and I think that that probably would have been the case.

FROM THE AUDIENCE: And I think that was a question actually raised by one of the comments that Dr. Thorpe made, that in some cases, you can actually see less and what would you do about it. So that was why I raised it.

DR. QUARMBY: Valerie Quarmby, Genentech. Two really wonderful presentations.

I'd like to just add a comment to something that came up earlier in the discussion. I believe Dr. Siegel and, also, Dr. Schellekens mentioned the concept of standardizing antidrug and

antibody assays, and I would just like to comment that many years ago now, a group of diabetologists around the world got together to try and standardize methodologies for anti-insulin antibodies to try and better understand which of those methods did not have clinical utility in predicting sequella.

It took five international workshops over a number of years to get to the point where the willing participants actually even understood what the key variables were to keep under control.

So I would just like to caution the audience that, in fact, I think these assays are really important, but to get those assays designed correctly and in appropriate agreement is not a trivial task.

DR. ROSENBERG: Thank you.

DR. SEAMON: Ken Seamon, Amgen. Again, I would also like to thank the presenters for excellent discussion and presentations.

The discussion so far and the questions have pretty much been focusing on comparing

immunogenicity between two different products, but when we consider how these are actually going to be used in the marketplace, I think there is the issue about switching and substitution that will inevitably come up.

The EMEA, in their draft guidelines, have recognized this by putting in specific references to unique labeling, as well as banking serum, importantly, to maintain the physician to be able to monitor and make decisions regarding use.

My question to the panel is is it possible and can you comment on the possibility of having two different products, each of which have relatively low immunogenicity individually when studied in controlled trials?

How does one assess the potential for administering one product to a patient who has seen another one, another product, potentially showing some very significant adverse immunogenicity or clinical sequella, because that will be an issue? I think we already see that for cross-reactivity with the erythropoietins, even with insulins.

So I would be curious, your comments about how to manage that and, also, whether, Huub, you think your models might be predictive in any way of

that.

Thank you.

DR. SCHELLEKENS: I know this is an issue, for instance, in the beta interferons and there have been studies going and there are studies going on, switching patients from one beta interferon to another. As to results, as I know them, I think what you need at least is a washout period between the two different preparations.

DR. SEAMON: Is there a basis to make any predictions about the possibility of seeing enhanced immunogenicity due to preexisting antibodies that were not significant, but were, for example, exacerbated by administration of other products?

DR. SCHELLEKENS: I don't think you can answer the question, but theoretically, if you induced antibodies from a high immunogenic preparation, you could argue that even a low

immunogenic preparation will be able to be enough of a trigger to keep the antibody production going, and that is something that if you change at least in that interferon that has been looked at and it seems that you--what you need between the two treatments is a rest period for the patient to avoid that, that the antibody production keeps going on. But it is an issue.

DR. STARK: Yafit Stark, from TEVA Pharmaceutical. I would like to ask the question about the importance, the clinical importance of the neutralizing antibodies.

Currently, there are three different interferon betas in the market. All of them are producing neutralizing antibodies. It's a function of the dosing frequency, duration of treatment, and the various interferon types.

However, I haven't seen, until today, any guidelines coming from the medical practice what to do when a patient develops neutralizing antibodies. Should we switch over? Should we stop their treatment? So the question to you is what we, the

developer of the biopharmaceutical generic, should learn from the experience of the innovators about their importance.

DR. SCHELLEKENS: I think it is clear from the studies that high levels of neutralizing antibodies affect the clinical efficacy.

The problem is to come--and these studies have all been done on a population level. The problem is how to translate this to an individual decision of a physician whether it's worthwhile to keep treating the patient, and that is another issue.

I know in Europe a number of clinicians are putting together a very big trial to be able to answer that question, but that's a difficult issue, and maybe, in the end, that may be impossible.

Look at the alpha interferon, where you have an extensive database, extensive experience, and that's a simple disease is to look at the effect of the antibodies, because you can follow the lack of efficacy more or less on the level of the liver enzymes.

Even in that case, it has been impossible to put a predictive value on an antibody level for the outcome in an individual patient.

MR. LISK: Alan Lisk, Barr Duramed. Just a specific question to Robin.

You used some older data to prove a point, be it positive or negative, but were the tests validated and is the material that you used for your immunologic comparison highly characterized?

What I'm getting at is aren't we at risk at using old data, whether it proves our point or disproves our point, with all the variances in these kinds of assays?

The second point is kind of broad to the entire panel. I heard a lot of negativity, can't do this, can't do that. Is the logical extension to withdraw all biologics that are currently on the market due to lack of data? Can we have some light from the panel as to how we should move forward on these things?

DR. THORPE: When I said it was old data, I mean, it's not that old. It's about six years

old.

But I think the point to make with that study, which may have relevance for other people doing these kind of studies, is that we set those assays up in our lab and we used the same assays for everything.

So we had complete comparability of the assay design and even the people who did them were the same people.

So if you compare what you find with two different products, you are using the same assays. So I think that is the genuine comparator.

If you went to different assays or even assays done in different labs, you would have a real problem. The assays were validated for the purposes that we were using them, but that's not necessarily all encompassing, if you see what I mean.

So that would be my only real response to that.

MR. LISK: I just worry, as the FDA has told us, you know, an invalid, unvalidated assay

gives us dubious results, even if it confirms our theories.

DR. THORPE: Yes. You would have to validate them and you would have to validate them for the purpose that you're going to use them. That's the message, I think, every time.

MR. LISK: No light at the end of the tunnel?

DR. ROSENBERG: There's always light at the end of the tunnel.

FROM THE AUDIENCE: I have a question for the panel on generally doing immunogenicity studies. If a manufacturer of a follow-on pharmaceutical protein, and one of the many products that exists on the market where you have immunogenicity or, you could say, antibody responses in patients in the five, ten, twenty percent range, clearly, one can do clinical studies then to demonstrate whether you are similar to the existing products on the market.

With the example that is often raised with EPO, since the existing manufacturer's rate of

immunogenicity is so low, clearly, a follow-on protein manufacturer can't do clinical studies to demonstrate that they're at that low level.

However, it seems to me there is still, I guess, a question to the panel, is there still not some value to having some immunogenicity study done with, say, a follow-on EPO or a GCSF to demonstrate at least that the level is below some very high level; for example, it could be one in ten, one in a 100 or whatever type of thing.

So even though you can't get to that low level in clinical studies, you can avoid a serious risk.

DR. THORPE: I think you're absolutely right. I think it certainly is the case that some biologicals seem to have or at least the potential to have an alarming rate of immunogenicity. Others don't.

I think GMCSF is a glaring example. It is very easy to find antibodies induced with GMCSF. But I'm sure that if we had carried out the kind of studies we did with GMCSF with EPO, we would have

never found any antibodies.

You're absolutely right. You have to--I think it is worth having an initial look if you find this kind of real glaring example of immunogenicity, you need to consider what the consequences of that are.

But you can't assume, just because you don't see something, these limited studies, that you're not going to see something in the fullness of time, if you like.

DR. ROSENBERG: I think more discussion will focus on that in the breakout sessions, because clearly one has to have a certain level of comfort and where that comfort level lies pre-marketing versus post-marketing is, I think, an important thing to discuss and perhaps an interesting ethical issue, as well.

One more question.

FROM THE AUDIENCE: I would like to have your opinion about the relevance of long-term tox studies performed with high dose proteins in animals from your immunologic point of view,

especially if you consider, if you give high doses into animals at a long period, I think it is clear that they will show immune responses and they will show organic lesions indicative for protein overload, like kidney lesions, for instance.

So I would like to have your opinion about the general relevance of these kind of studies in animals from the immunogenic point of view.

DR. THORPE: So you're talking about animal studies using human proteins.

FROM THE AUDIENCE: Yes. Correct.

DR. THORPE: I wouldn't think they're very informative, in most cases. I would only think that they are useful in these kind of comparative studies, like I showed that mouse version. I can't really see much use of animal studies for immunogenicity assessment apart from that.

DR. ROSENBERG: We will finish it up with Paul's question.

DR. CHAMBERLAIN: Thanks, Amy. Just a follow-on from those two previous questioners.

I guess really for molecules like GCSF and

for EPO, the common human forms thereof, the dog is quite--the dog does produce a B cell response and a measurable immune response, humanly, within four to six weeks, and that does develop, in both cases, into clinical sequella which could be reminiscent of the human disease process.

So an erythroid hypoplasia appears with a relatively high incidence in dogs dosed for up to 12 weeks with recombinant human EPO and that is definitely cross-reactive, neutralizing to the K9 growth factor, and, indeed, the same is true in dogs with GCSF. If you dose them for long enough, they will become thrombocytopenic, indicating that you have a neutralizing response.

So I kind of wasn't sure about the point that Huub made earlier on about the non-discriminatory value of preclinical models when it comes to relatively subtle structural alterations.

I'm not sure if there is enough data there to actually make that point with confidence.

DR. SCHELLEKENS: But you agree with me

that the dogs didn't predict what was going to happen in the human patients. I mean, they showed you what you would expect if the patient would make antibodies, but they were in no way predictive.

DR. CHAMBERLAIN: Well, it depends on how much you look. I think the TEPO story shows that you can map epitopes which could have predictive clinical problems, and Amy has always said, well, are we looking hard enough in the preclinical scenario, and that is the question I would like to put up.

I don't think the preclinical database is strong enough to really discount the value in a comparative scenario.

DR. SCHELLEKENS: But how would have the fact that you know the T cell epitope would have predicted what is going to happen in patients?

DR. CHAMBERLAIN: It is useful as a comparability tool in the absence of being able to power a clinical study effectively to compare the relative immunogenicity of an originator with a follow-on product.

DR. ROSENBERG: I could see where it might not have predictive value, but it might be useful as a comparative, because if it takes four to six

weeks for 80 percent of the dogs to develop neutropenia or anemia and you have a different EPO preparation, let's say, that comes along and produces it in two weeks, that might raise a red flag, or if that EPO preparation doesn't induce it at all, that might tell you that there are some differences.

But certainly in terms of prediction, I think that is correct. I don't think that they are helpful that way, but they could be helpful in other ways.

Well, thank you very much, all. This was an unexpected gift to have a chance to question the panelists, and I thank you all for your attention and good questions.

[Applause.]

[Recess.]

DR. ORLOFF: People come and take their seats, please. Let's get started. Welcome,

everybody, to Session No. 6, which is Approaches to Clinical Safety and Efficacy Studies.

I'm David Orloff. I'm the Director of the Division of Metabolic and Endocrine Drugs in CEDR, and I will be the moderator for this session. I know Dr. Siegel's talk is about a half an hour long. I don't know how long Dr. Ben-Maimon's talk is.

We will reserve questions for the end, if there is time, in this hour and a quarter session.

As people know, the format here or the game plan is that these sessions set the stage for the breakout sessions that will occur this afternoon in repeated sessions, starting at, I guess, 1:30.

With that, let me introduce Dr. Jay P. Siegel to give the first talk. He is President, Research and Development, of Centocor, Inc., and the title of his talk, in keeping with his instructions, is the very imaginative approaches to clinical safety and efficacy studies.

Dr. Siegel.

[Applause.]

DR. SIEGEL: Good morning. A few comments before I move formally into my talk.

One is that any opinions expressed are my opinions, not those of anybody else in particular, not cleared by anybody else in particular, although many of the facts and concepts I am greatly indebted to a large number of colleagues over many years.

Those of you following my slides will find that there has been a little bit of editing and shuffling, so maybe 80 or 90 percent of what is there is the same, but it's a little bit different.

The other thing I want to point out is I will, and I will try to keep this down, but I will, to some extent, go back over some of the issues that you heard discussed in sessions on analytics, PK/PD, and immunogenicity.

It is really very hard, at least for me, to separate out these issues. All of those areas matter a great deal, but in this setting, they largely matter in determining the extent to which

you do or do not need to do additional clinical trials. So it is very hard to talk about the factors involved and the nature and design of clinical studies without going back over some of the areas already discussed.

So I want to start with a few background comments and just to set the stage and concepts, as well, and then move into a discussion of the factors that determine the types and size of clinical trials that one might want to do.

So I put on this slide what I consider a fundamental question of this talk; to some extent, one of the more fundamental questions of the whole debate here. To what extent can non-clinical information about similarity of a protein product to an approved product diminish the amount of clinical data required to make determinations regarding safety, efficacy and equivalence?

I suppose that since clinical costs are the largest share of costs for development of any given product, at least innovator product, this is perhaps the most critical issue regarding whether

follow-on products might be able to be developed in a more cost-efficient manner.

Now, I have stated here a fact, and I know there will be a lot of disagreement around this, I believe it's a fact. I think we will all disagree to a large extent, and I don't want to deny that, over the extent that it applies about--so the statement is no amount of non-clinical testing of a protein product can assure it will have identical effects to another product.

I think we have heard some healthy debate about how much assurance we can get in different settings and I'm not going there right now, just to say there is always going to be some clinical effects that one cannot predict from non-clinical testing, and, therefore, a risk of inferior safety and/or efficacy will always remain, and I would argue, in some cases, perhaps not all, this is substantial.

Of course, it is clinical testing, the subject of this talk, that can then further limit that risk. So I really want to set the stage for

this talk in terms of approaches to risk reduction.

In that regard, it is important to keep in mind that clinical trials don't make exact determinations of an effect or even exact determinations of a comparison of effects between two products, if you are comparing them, but rather they estimate effects. They give ranges and estimates and the precision of those estimates or the narrowness of those ranges, if you will, increases with the size and the number of clinical trials.

The more the data, the more precise, you know whether there is a difference, whether there isn't, or really what you're dealing with in clinical is not there is or not there isn't a difference, but that you can say, within a certain range, that things are highly similar. The more studies and the larger the studies, the narrower that range becomes in terms of outcome data.

So it follows that less clinical testing of a follow-on product means lower assurance about equivalence of effects and greater risk. More

testing, of course, means higher assurance and lower risk.

I pulled this off on a separate slide to highlight, for a couple of reasons. The risk of inferior efficacy or safety is not balanced by any significant possibility of superiority. Perhaps from what we just heard, it's possible that a generic might turn out to be less immunogenic.

For the most part, innovator products are optimized by dose for certain effect and generic products are trying to be the same. So we're not holding out much hope of superiority. We are holding out hope of equivalence.

The reason I highlight this is two. One is that what I have just said, that more data diminishes risk, the more you have, the less risk, applies to all clinical development.

For an innovator product, at the time you approve it, there is always a risk. You only need to read the papers to realize the risks of learning about adverse events that haven't been found in a pre-marketing period.

In that setting and in most settings outside the generic area that the FDA operates, the way it operates is it balances those risks against

benefits. So a product is providing some new medical advances, usually addressing medical needs, and there's only a number of FDA policies that specifically allow for higher risks or higher uncertainty in the face of more benefits and addressing unmedical needs, the E-1 document on the amount of safety data needed mentions this, the accelerated approval mentions--applies to products that address unmet medical needs in serious illnesses and allows more uncertainty about certain of the effects at the time of approval.

But we're in a different paradigm here where we are not balancing those risks against that type of potential benefit.

Another reason to just note the difference here that it's not balanced out benefit is that when we think about clinical trials, we do need to think about feasibility, the ability to enroll those trials and the ability to get informed

consent.

So as we think, in some of these settings, we may be thinking about asking patients to enter a trial where they would be told, as part of informed consent, they can either have this approved and marketed drug that has been well studied, or be randomized to get another drug that we hope is as good, but isn't likely to be better and might be inferior, or, alternatively, we would tell them that they could just choose to get the first drug and not enter the trial, and sometimes those trials are hard to conduct.

So coming back to the fundamental question that I asked up front, which is in the first bullet here, shortened.

To what extent might similarity in nonclinical studies reduce the need for clinical studies? I am going to address this or break this down, based on the comments I just made, into three questions.

How much risk? So if there is going to be some risk, at the end of any study, nonclinical or

clinical, how much risk of inferior clinical safety and efficacy is acceptable in the approval of products that offer no new benefits in safety or efficacy?

What is the nature of the residual risk of clinical inferiority after nonclinical testing has been completed? So what do we know and what don't we know about the possible clinical effects and how can we best address that risk in clinical studies?

So as to the first question, and most of my slides will deal with the last two, how much risk of inferiority is acceptable with no clinical benefit? Well, I think we all agree that safety is critically important, and, therefore, this is very small. Some might argue none. Some might--if there is none, that, I think, is a problem. I don't think it is consistent with the way we approach a lot of new drugs.

Any new drug has a certain amount of risk compared to a drug that has been on the market for many years.

That will be true, by the way, if we

require follow-on products, protein products, to do all the testing that we require an innovator to do, it will still not have the years of experience on the market and information that the innovator has.

So there will be some risk. If we say no risk is acceptable, obviously, we have a highly restrictive situation.

Perhaps the answer should be, and that is why we are here discussing it, that some amount of risk is available because although there may not be clinical superiority, there are other societal benefits to having follow-on proteins.

So what then are the factors influencing the amount of clinical data needed? I have divided them and I think they can be thought of as falling into three general areas.

One is the nature of the comparability claim sought. In the handouts you have, it may say similarity instead of comparability, but when Keith opened the meeting and he defined similarity for us as meaning the same clinical efficacy and safety effects, I figured that wasn't what I meant. So

I'll come back to that just momentarily.

But what is the nature of the claim sought? Factors about the product and then factors about the indications, and you could call this a fourth group, the effects of the product in that indication and elsewhere.

So, of course, you can imagine a situation where a follow-on protein product might seek many different sorts of claims. One might be just, well, we're safe and effective, we're much like this product that is already on the market, and perhaps it takes a little less or a lot less clinical data to establish it is safe and effective because of the similarity.

But this might be a claim that it is safe and effective, but not one that says that its effects are identical or highly similar.

You move to areas, and I don't know the terminology of the generics world as intimately as many others in the audience do, but you move into--so I may misspeak, but you move into areas such as therapeutic equivalence to show that a drug

is not only safe and effective, but actually has efficacy on the same order of another drug, is a much more complex thing to do clinically.

So if you are unable to presume this from nonclinical data, this is a major burden clinically.

For example, if you have, and this is taken from a clinical setting, not a typical, if you have a disease in which five percent of patients on placebo will spontaneously improve, and 50 percent on a drug, and now you come along with a follow-on drug that has the same 50 percent response rate, to show that drug to be significantly better than placebo in a head-to-head trial would take about 50 patients.

To show that it has an effect that is within ten percent of the drug that is already out there, let's say, has a 50 percent response rate, you want to make sure you are at least within ten percent, that would take 850 patients. To show you are within five percent would take 3,400 patients.

So to actually show equivalence clinically

is a very difficult thing to do.

And interchangeability, a claim that would allow patients to switch back and forth, pharmaceutical, again, I'm not an expert on this, but one drug to be substituted for another would raise additional issues regarding the ability to follow patients for adverse events that may arise in the post-marketing setting, and is something that would, therefore, require a lot of assurance that those events will not arise in the post-marketing setting.

Now, moving on to issues about the product, we start with what has been the focus of most of yesterday, the extent of nonclinical characterization of the product, and I note here, also, and any differences found.

This has been a theme that I have heard come up at this meeting a couple of times and I think it is an area where we need to be very clear about it.

Some people conceptualize this situation as, well, we have this tremendous analytical

characterization and we'll analyze everything and if we see no differences, we'll know we're okay.

Reality is the nature of these products and their complexity and their mixtures is that if you can analyze everything, you will find differences in most, if not all cases, and really a lot of this debate is about when you find differences, how do you determine which differences make a difference clinically and which ones do not.

Part of getting at that is understanding the structure-function relationship and, certainly, the better you understand structure-function relationship of a molecule, the better ground you are on in moving forward in this regard.

Minor variations in structure are likely, as I said. Typically, knowing the structure-function relationship is something you know. An innovator often has information based on their clinical trials and see how their product has performed through trials and in the marketplace, with some amount of variation in production. Those data are not necessarily available to a follow-on

protein product manufacturer or likely wouldn't be under current law.

It is also important, I think, to note that different functions may result from different structural features, particularly true in these large and complex molecules.

As a common occurrence, it can result from their interacting with different receptors or interacting differently with the same receptor. Some parts of the molecule may bind, others may trigger internalization or other activation.

They may bind in different tissues and penetrate to those different tissues differently. So features of the molecule may affect pharmacokinetics or the biodistribution. Some may affect the immunogenicity and antibodies. You have parts that affect the binding, parts that affect effector functions, effect PK and so forth.

Being similar in some areas does not guarantee being similar in other areas. This has important issues, I think, in terms of clinical testing, because there's a lot of discussion; for

example, if you show equivalence in one indication, does that tell you that you are equivalent in another indication, or if you show equivalence on one outcome, does that tell you you have equivalence in another outcome.

Well, there might be some inference that one can make in some settings, but one needs to be cautious about that. One can see different outcomes resulting from different changes in a molecule or from different patient populations that are apparent in some indications and not other indications.

So you can have two monoclonal antibodies that bind a target the same way. Well, you have a number of anti-TNF therapies on the marketplace, three of them, and in some indications and in some adverse events, they appear identical. In other indications and in other adverse events, there are suggestions that they are different. Obviously, they have physicochemical differences.

But it points out that they have some activities that look quite similar and some that

look quite different.

Speaking of differences in different clinical settings, if you were to do an equivalence of a--if you had Eprax on the market, since we have talked a lot about PRCA, it is probably a worthwhile example, and you were to bring along another erythropoietin and do extensive studies of it, it would take hundreds of thousands or tens of thousands to really get a good estimate of PRCA incidence.

But if you were to do that and you had a product that had an elevated PRCA incidence, like Eprax had for a while, but you were to do those studies in cancer patients, you wouldn't detect that.

PRCA does not appear, despite extensive use of Eprax in cancer patients. It appears in the renal failure patients.

So finding equivalence. So Eprax is equivalent to the Eprax before the changes that caused increased immunogenicity, if you study it in cancer. It's not not equivalent of you study it in

renal failure.

So as I say, there are implications that equivalence in one setting does not necessarily imply equivalence in another.

I'm not going to comment at great length about micro-aggregation. It has been discussed, particularly with regard to its relationship to immunogenicity.

It influences a lot of clinical aspects of a protein where you are concerned about micro-aggregation, where a product exists, as some do, as micro-aggregates. I would suggest you will want a significant amount of clinical testing, because they are important. They are not always easy to measure, because they are influenced by the assays that measure them.

I think I also have a couple slides here about immunogenicity, a topic that I will go light on in this point, because we have had an excellent discussion of it this morning and we will continue to do so in the future.

Only to note that it is difficult, as we

have heard, to predict short of doing the studies and that the likelihood of immunogenicity is likely to be a factor, based on the--you can try to predict based on the peritoneal molecule, but, of course, there could be, in a follow-on molecule, some changes that change immunogenicity, but there are plenty of examples of that, in addition, of course, to the likelihood of immunogenicity as the implications of immunogenicity, and we know there are some settings, particularly as discussed, where neutralization of endogenous homologs is going to be of particular concern where one would argue, I think, for a higher level of risk reduction through clinical studies.

I want to speak a little bit about this concept of inactive ingredients, and I have to confess to being not, also, highly familiar with the formal definitions of process, product-related impurities and active and variants and so forth, but I would like to say that in the case of biologics, at least, it is very important not to ignore the minor contaminants of a product, whether

or not they are unrelated to the product in terms of structure and function, such as these small amounts of rubber leachates from the stopper in Eprax that appear to have a role in its immunogenicity, or whether they are related to the product.

Very small amounts cannot be ignorable. Very small amounts of material can increase immunogenicity. Very small amounts of product variance can be bioactive in important ways.

I have on the slide a number of types of variants that are commonly observed in biological molecules. I think an excellent example of--I don't have the numbers for it, because I was just thinking about it yesterday, but an excellent example of a small contaminant related to a product that can have a huge effect occurred in the development of one of our products, Reopro.

Reopro is an FAB fragment of an antibody that binds platelets and inhibits their aggregation, and it is manufactured by manufacturing a whole antibody and cleaving the FC

off to leave the FAB.

It turns out that very small amounts of residual FAB that can occur and did occur in a lot during early manufacturing, pre-licensing for clinical materials, I think amounts that were on the range of one in a 1,000, there are probably folks in the audience that could correct me, while having the same in vitro effects in terms of blocking platelet aggregation when used in vivo caused clearance of platelets and significant thrombocytopenia. So this is a one part in a 1,000 contaminant.

So aggregation. We have talked about degradation products, too, because in immunogenicity, it is important to pay close attention to small contaminants, and, of course, they can have their own direct toxicity, but they cal also importantly interact in various ways with the active ingredient.

There was a product I mentioned in one of the breakout sessions yesterday of a small amount of metalloproteinase, which was no problem when the

product was in a vial, but when it was put in a syringe, the zinc in the needle of the syringe activated the metalloproteinase, which degraded the product.

Now, to the third and last factor I wanted to speak about was the factor of the indication, the clinical indication and its effects and the effects of the product and how those might influence the necessary clinical testing.

So one that has been mentioned that I think is quite important is the steepness of the dose response curve, where you may have some uncertainty about a different potency or higher or lower degree of efficacy.

When you have a steep dose response curve, then a little bit more or a little bit less can make a big difference. That is going to be a bigger concern and certainly a need to ensure that you don't have problems in the clinical trials.

That then relates, also, to the next bullet, the nature of the indication and the proven effects of the innovator, both in terms of how

serious it would be to inadequately treat the disease or, in other words, the implications of under-treating or over-treating.

So if the therapeutic window is narrow and the likelihood and implications of under and over-treatment are high, then that is an area where you are going to want to do some substantial clinical studies to ensure that you are on target.

And under-treatment, of course, can occur because of a variety of reasons; if a product is less potent or has a slightly different PK or enters tissues differently, and will be less important in certain diseases, where you are dealing with symptom relief, or perhaps in chronic diseases, where you can get feedback and if you are under-treating the patient, you can bump the dose up, escalate the dose, switch over to a different medication.

I think, as noted in the second bullet, however, that we should exercise extreme caution in accepting the risks of diminished efficacy with those therapies that reduce serious morbidity or

mortality and prevent irreversible damage.

So to take a therapy used in the setting of acute MI or acute stroke that reduces mortality and morbidity and study a therapy that may be inferior is going to be a thing only done with extreme caution, if done at all.

And I would note that a lot of this applies to quite a few biological therapies.

Over-treatment also can occur for a variety of reasons and there was a classic, a well known, that could be picked up by analytics certainly today, but a story regarding changes in the manufacture of TPA by Genentech that led to a longer half-life and required significant dose reduction in a setting where over-treatment could lead to intercranial hemorrhage and other serious complications.

Now, also important in terms of what you are going to need from clinical trials is the ability of the trials in a given setting to measure the important effects. That is going to be a critical determinant, as it is in any development

program of the size and nature of trials.

It will be easier to exclude differences where trials of modest size and duration are able to measure a drug's effect, and we heard discussion of this in the immunogenicity setting, talking about the easier to compare the immunogenicity of products that have maybe a ten or five or twenty percent immunogenicity than of those where it might be one in 10,000 patients.

The timing of events, of course, is a critical one. Where there are markers or outcomes that occur early, at a high enough frequency, I think there's better possibility to get assurances from clinical data that you have an acceptable product.

I think the existence of useful markers--well, let me go into the next slide, since I have started talking about that. The existence of markers and surrogates is an area that hasn't been explored as well as it might be in terms of its ability to assist the development to follow-on protein products.

I have heard it said by some that very few are validated; therefore, very few can be used. It is true, there is no question, that surrogates and

markers can often be misleading, but they have--the meaning of validation and the meaning of a surrogate is different in different clinical settings.

It is one thing to have a surrogate for an efficacy outcome and say I have a completely new therapy in a total new class of drugs that might have a different mechanism of action than, say, one example I've used, we were talking about glucose response in patients on insulin.

I might have a totally new therapy that impacts glucose utilization or that might impact glucose production. That might be an enzyme that breaks down glucose or something that cause diuresis of glucose, and it might have the same effect as insulin on glucose levels, but I would hardly say that that proves that it is an effective treatment.

On the other hand, if I make an insulin

that is by analytics as close to identical to another insulin as we can determine and then it has the same effects on glucose, that tells me a lot about the effects of that drug.

If I'm somewhere in between, it's a variant of insulin, may have a different half-life, whatever, it has an intermediate usefulness.

So I think in follow-on products, there are more surrogates that may be useful in predicting clinical outcomes.

There are two limitations, however, to the usefulness of such surrogates, even if one accepts one can use unvalidated surrogates.

One is that the data that show that those surrogates are useful are the data that come from the innovator. So you have a body of data from the innovator showing, well, my drug affected this surrogate and it affects this long-term outcome and, therefore, I know they are related. So I guess there are some legal questions about whether such data could be used by a follow-on product.

Another caution about that would be that a

surrogate, even if it is acceptably valid, is a surrogate for a certain outcome and as I have talked about, these molecules are complex molecules. Different parts may affect different outcomes.

Having the same efficacy on a given outcome does not mean the same safety profile, does mean the same efficacy in a different disease or even on different mechanisms within the same disease.

I think I'm in my last few minutes. This is something I have highlighted already, that effects of a drug may differ in various indications. I spoke before about the fact that they may involve different functional aspects of the molecule, but there are other reasons, because of the nature of the disease, the different tissues of activities, the different concomitant medications that equivalence in one indication may not imply equivalence in other indications.

So a few remarks in summary. One not covered on the slide is that--but I think that

really is brought out by my talk, that the approach to this problem should be one of risk reduction. How do we manage the risk? How do we reduce the risk?

In doing that, it is important to note that it really is a multidimensional problem. In the introduction to these proceedings, we talked about proceeding from simple to complex, but even in that introduction there was an acknowledgment that we are not simply talking about simple molecules versus complex molecules.

There is a whole bunch of dimensions that you can measure the risks, the risk of immunogenicity, the risk of inadequate therapy, the risk of product variance that could be active and so forth, and each of those risks needs to be looked at separately, I think, in terms of what its implications are, in terms of what clinical trials are going to be needed.

In concluding, then, nonclinical testing can reduce, but not eliminate the risk that a follow-on product is inferior. The more the

clinical testing, the more risk can be reduced. It is going to be proportional. It is never an absolute amount that eliminates all risk.

And specific aspects of the similarity claim and use being sought, the product and its characterization and its indications and effects will impact the amount and types of clinical data needed to reduce risk.

Thank you.

[Applause.]

DR. ORLOFF: Let's see if I get this right. Our next speaker is Dr. Carole Ben-Maimon, who is President and Chief Operating Officer of Duramed Research. Excuse me for mispronouncing it.

She has another imaginative title, also, following the title of the session, and I imagine Dr. Siegel's talk gave us some food for thought for the breakout sessions and this one will, too.

Dr. Maimon, come on up.

DR. BEN-MAIMON: Good morning, everybody. I just corrected the disclosure statement which says I'm not affiliated. I can assure you that is

not the case. I am very affiliated.

I work for Duramed Research, which is a wholly-owned subsidiary of Barr Pharmaceuticals, and I manage the innovative portion of our development program.

So I actually work on the innovative side of the business now, conducting clinical trials for our innovative products. My past life, I actually worked in the generic side of the business and was Chairman of the Generic Pharmaceutical Association for several years.

So I think when I look at all of these products and I look at these discussions, I realize the value to the patient in having both access to new and novel therapies, as well as access to generic drug products.

Dr. Siegel and I are much in agreement, quite honestly, on many of the things that he discussed this morning; in particular, the benefit-risk assessment and how important it is to keep in mind benefits and risks as we look at developing new drug products and approving generic

drug products, and, hopefully, someday, generic biopharmaceuticals

We should never forget that generic drug products have really benefitted in a huge way the patient population. They control costs. That doesn't benefit necessarily the individual patient, on its face, but when you look inside, it actually does benefit the patient and it benefits the individual patient, because if the patient can't afford access to a drug product, they can't take it the way it is prescribed, they can't reap the benefits from that drug product.

So controlling costs, although, on its face, may seem like a completely nonscientific, nonmedical objective, it isn't. It is actually a very important medical objective.

The other way the generic drug products have benefitted people is they have stimulated innovation. As we challenge and compete with each other on the generic level, companies are challenged to innovate and develop new products.

So I do think that there are significant

benefits that have been reaped from generic drug products and will continue to be significant benefits as we move into the future, and I hope that generic biopharmaceuticals will make the same contribution to the patient population and to medical care as we move forward.

I wanted to start with a little discussion of what the objective is. One of the things I think we ought to say is what the objective is not. We are not talking about an abbreviated chemistry manufacturing and control section.

The generic industry understands very well that these are difficult products to manufacture, that the processes and the manufacturing processes have to be validated, they have to be robust, they have to be reproducible. Products have to be made in a consistent manner in order to ensure their quality, and, clearly, we would apply and do apply those same standards to our generic drug products and would apply them to generic biopharmaceuticals.

It is important to note that in many places around the world, these products are

actually already being made, already being sold, and already being used by patients. There are biopharmaceuticals on the market that are comparable to products that we have today in the United States.

The studies to actually prove it have not necessarily been performed, because here we are today discussing what those studies should be.

But, clearly, products made in other parts of the world have been demonstrated to achieve the same outcomes and have the same safety and efficacy profiles.

Again, I would reiterate that we are not looking for an abbreviated process for the CMC section. What we are looking for is an abbreviated process for pharm/tox and clinical development, and really that abbreviated process should be based on these four points, as we see it.

It should be based on sound scientific rationale, at no way should put at any risk--I shouldn't say any risk--all risks should be benefit-risk associated and should be calculated

and well controlled in order to ensure safety and efficacy, comparable safety and efficacy.

We should not be performing unnecessary clinical trials for the sake of re-proving things that have already been proven. They may not have been known ten years ago when the brand product came out, but if they are known today, there is no reason to reprove what is already understood.

The regulations should allow for interchangeability, not necessarily require interchangeability.

There should be an abbreviated process that would allow for what I will call "me too" type products, products that have the same safety or comparable safety and efficacy profile, but may not necessarily be interchangeable, but, again, the process should allow for the assessment of interchangeability when appropriate.

I'm not going to focus on immunogenicity. We just spent what I think was a productive morning listening to two experts who clearly highlighted a lot of issue that we need to deal with as we go

forward on immunogenicity. So I'm going to avoid talking about that and focus on clinical safety and efficacy.

And when I talk about safety, I know it's a little bit hard, and efficacy, for that matter, to isolate that from immunogenicity, but there are safety issues associated with all drug products, as we know, from benefit-risk assessment, that are not necessarily interrelated to immunogenicity.

It is our position that clinical trials for some products will be unnecessary. When products are well characterized, not glycosylated, and easy to identify and well understood, have been on the market for many years, there may even be multiple manufacturers, there may be no necessity to do clinical, full-fledged clinical testing.

Clearly, PK will be done more commonly than just clinical trials and PK/PD, when necessary, should also be performed.

I wanted to sort of--I want to see if I have another slide. I don't.

I want to sort of refer you back to one of

the presentations yesterday, and there was a slide that was basically from top to bottom with two arrows.

One of the things that came to my mind as we were talking in the breakout sessions yesterday and listening to the plenary sessions, I was actually very reassured by the fact that the analytical methods that were presented yesterday and the collection of those analytical methods was able to identify so many differences with regard to the chemical composition and the molecular structure of some of these products.

Again, I think Dr. Siegel and I agree that it is very difficult. The challenge is not to identify the differences. The analytical methods, the characterization exists in order to be able to demonstrate the differences.

The question we really have to answer is are those differences of any clinical relevance, and that is actually, in some ways, a more challenging question than are there differences.

With history and with experience and

having used these products through time, the things that will actually modify, the characteristics that will actually modify the effect and the safety profiles of these products can be identified.

So I think that as you move from the process of PK animal studies and up to clinical testing, although you lose sensitivity, to some extent, you do gain the ability to look at clinical relevance.

So for simple products, as we talked about before, you may not have to do as much testing. Products that are well characterized will describe where the primary, secondary, and tertiary structure can be identified, where we know what's happened, they have been on the market.

Chemical characterization, possibly PK studies, and ultimately maybe PD can be performed. When you need to really assess whether there is a difference, whether that difference does translate into a clinically relevant difference, then more studies in animals and clinical, full-fledged clinical studies may be necessary.

There are clearly some products that cannot be assessed through pharmacokinetics, products where there is no assay, where they act

locally, some topical products, and there clinical studies would be appropriate to assess the therapeutic equivalence of those products and to evaluate the clinical impact of those products, but those products tend to be many fewer and most products you can perform analytical studies to do PK and to look at surrogate markers and pharmacodynamic measures.

I'm actually going to start at the bottom of this slide. This slide speaks to the design of these clinical trials. Clearly, we don't want to be doing studies, as I said earlier, to re-create the wheel.

Everybody knows that erythropoietin raises hemoglobin. That doesn't have to be proven. Quite honestly, today, which we didn't know 15 years ago, raising hemoglobin does have a positive effect on morbidity.

That was not necessarily known. When I

was a fellow, we actually did some of the clinical trials on some of the products when I was studying, and it was apparent we had to actually show that not only did we raise hemoglobin in patients who were in dialysis, but that that rise in hemoglobin actually translated into a clinically reasonable, clinically important effect.

Today we know that. We don't need to show that. We just need to show that these products are comparable in their effect on hemoglobin, and I will speak to that issue as I move further when we talk about surrogate markers. But the extensive data that has been generated on many of these products, as they have been on the market for years, should be used and should be taken into account as we design these development programs.

In addition, PK and sometimes PD should be the crux of our program. If we can do PK and PD studies and demonstrate equivalence, we should rely on those. There is no need to go further into full-fledged clinical testing, and when we do design the trials, we should target the trials to

answer the questions that are being asked by the differences that we're seeing either in the PK profile or in the characterization of the product.

What I mean by that is if there is a specific difference that we think may translate into a neutralizing antibody or may translate into a specific outcome that may be different, that is what we should be evaluating.

We don't need to evaluate the whole host of things that were identified by the brand.

Surrogate markers. It's been sort of a tag line in the conversations and our discussions over the last two days. Surrogate markers are something we live with every day. We don't even think of them anymore as surrogate markers, but what is hypertension?

Hypertension is nothing more than a surrogate marker for cardiovascular disease. Products today, antihypertensives, show effects, long-term studies do show effects, but primarily what do you look at? You look at hypertension. You look at whether blood pressure comes down or

not, glucose, and eve plasma levels.

Plasma levels are merely surrogate markers for the rate and extent of exposure and the extent at which a product gets into the target organ.

So we use surrogate markers on a daily basis to help us make decisions in drug development programs. We should continue to use those surrogate markers. We should develop as many as we can. It saves patients a lot of risk and a lot of time as we enroll them in clinical trials, and we should continue to make use of those surrogate markers to help expedite and minimize the risks to patients.

The last comment I would like to make about clinical trials is the issue of ethics. It is very well recognized in the Declaration of Helsinki and in other regulations that clinical trials inherently have their own risks.

For those of us who design them, we may do an x-ray that isn't necessarily needed from a clinical perspective, but we need in order to evaluate the outcome or ensure the safety of the

patient.

We may draw blood. There are a whole host of procedures that patients become exposed to in a clinical trial setting that they may or may not be exposed to in a clinical setting.

So limiting the number of clinical trials and the animal testing that we do is really essential and really very consistent with regulatory policy throughout the world today.

A couple comments on safety. I was actually a little bit surprised at the breakout session yesterday, because somebody suggested that we do clinical trials, and they used this word, "to eliminate the risks associated with drug products."

That is completely flawed, as we all know, I think. Drug products inherently have risks. Any time you take a product, and as we have seen in the news, any time you take a product, there is some risk associated with it, whether it's Tylenol or some more complex complicated biologic product.

So it is inherent that we understand that clinical trials describe the risks. They describe

the risks in order that the patient and the physician can come to a decision as to whether or not the benefits of that product outweigh the risks.

So as we design our development programs for biopharmaceuticals, generic biopharmaceuticals, we need to keep in mind that we will never be able to eliminate risk.

What we need to do is minimize risk and recognize benefit and still look at that same benefit-risk assessment that we make every time we submit an application or the agency approves or denies a product what that benefit-risk assessment is.

Clearly, trials conducted by generic companies for generic biopharmaceuticals will be charged with the same record-keeping, the same assessments of safety, as any other trial that is performed by any other company.

Post-approval. We continue to do pharmacovigilance. We will continue to do pharmacovigilance. In certain situations, we may

need to do more, we may need to do less, but we are held to exactly the same standards of reporting, of annual reports, as every other part of the industry, and, of course, we will continue to do that and take tremendous pride in the fact that we do it and we do it with real integrity.

Finally, I guess, in conclusion, I would just like to say that it is very important that we recognize the advantages and the essential need for generic biopharmaceuticals. The generic industry is absolutely committed to ensuring the most robust safety we can and to doing whatever we need to do to make sure that these products are comparable.

What we don't want to do is be held to higher standards than the brand industry is held to.

With that, I would thank you all and turn it back over to Dr. Orloff.

DR. ORLOFF: Thank you very much. We've got about 23 minutes, by my watch, and so I think we will take some questions from the audience.

If you wish to ask a question or make a

comment, please come to the microphone, state your name and your affiliation, so we can get it on tape.

DR. SEAMON: Ken Seamon, Amgen Corporation. I would like to ask a question to both Dr. Siegel and Dr. Ben-Maimon with regard to assessment of safety.

I was particularly struck by your comment toward the end of your talk, and I agree with you that you cannot eliminate risk. There will be risk of taking drugs. It just needs to be assessed so that the physician can make the appropriate prescribing decision in consultation with the patient.

I would be curious how that works with getting a rating of interchangeability or substitution that could take the physician out of the picture, and it seems to me that the safety data is somewhat unique to these types of products, even in some type of an abbreviated approval process.

DR. BEN-MAIMON: I think that is why we

would like to leave the door open both for interchangeable products and non-interchangeable products.

I think the fact of the matter is there will be products where interchangeability can be the rule, because the safety profiles are comparable, and there will be products that are more complicated today and we may not, in today's world, with today's technologies, be able to demonstrate interchangeability.

So those products, we need an abbreviated pathway that takes that into account, as well.

But I do think that we need to make sure that we leave the door open to interchangeability for those products where it can be demonstrated.

DR. SEAMON: And could I ask either or both of you to comment on the ability of demonstrating that equivalence of risk or safety using pre-approval trials?

DR. SIEGEL: I think, in a sense, we are talking about two types of risks here; the risks that are associated with the drug, its safety

profile, and although they're the same risk, one can conceptually think also of the risks that the second drug differs from the first drug with regard to those or other safety concerns.

I think that to establish, to the point of Dr. Maimon's answer and your second question, to establish comparability of those risks in the preapproval setting, you can for more common risks, and for less common risks if you have substantial data.

If an event occurs, an uncommon event occurs, the variance around the estimate of its rate is about the square root of the number of times you see it.

So if you get a 100 cases of something, your standard deviation is about ten cases or ten percent and your confidence is plus or minus 20 percent.

So if you do a large enough study, to where you're seeing a 100 or a few hundred cases of the important adverse events, you can be pretty confident they are the same.

Otherwise, you have to draw inference from other than clinical data, and that, I think, runs risks or concerns.

Then the risks of uncommon events, of events that may occur in different clinical settings, events that occur in patient populations or with concomitant medications that are not studied in clinical trials, those are very real for innovator or follow-on products, and that's why we don't control all those risks with preclinical trials, with pre-licensing trials.

DR. BEN-MAIMON: I think some of your comments, the fact of the matter is that many of the risks and the less common risks are actually borne out with experience over time, and the generic biopharmaceutic actually benefits from the experience of the brand.

In addition, there are parallels that can be made across classes, which are clearly issues. We talk about the COTS-2s are a good example today, where we have class labeling.

So I think that for more common risks,

there is the ability to identify difference, as you said, and with less common risks, again, it's a benefit-risk assessment and you do it through pharmacovigilance, you do it through a whole host of other areas, but the fact that they are so uncommon often does make it less of a compelling issue to withhold the product.

The other comment I think I would like to make is we actually saw a case this morning, the Avonex case, where there was an improvement with the change. So there is nothing to preclude the labeling of a generic biopharmaceutical to say that there is less antibody formation, and if that is the result, there is an advantage.

DR. FIELDER: Paul Fielder, Genentech. We had a very nice discussion yesterday on the use of PK/PD and it was most of the consensus of the group that this would not be adequate to substitute for clinical safety and efficacy.

I would even bring in simple examples of growth hormone. We do not have a pharmacodynamic marker that would predict efficacy or safety.

Then I would like you to address one point. If it is unethical to do clinical trials with follow-on biologics, how is it then ethical to

treat children with untested drugs?

DR. BEN-MAIMON: Can I just make a couple of comments? First of all, with all due respect to everybody in the room, we have heard tossed around here all morning the word "consensus," and I don't think majority is necessarily synonymous with consensus. Clearly, the generic industry is a minority here and was at our breakout session.

But I would beg to differ with the comment that there was a consensus. There was probably a consensus within the innovator part of the industry, but not with the generic industry.

But with regard to the PK/PD issues, growth hormone is one product where it is very well characterized. It is not glycosylated. There are clearly parameters to measure growth hormone.

There are multiple products on the market. They may not be interchangeable, but there are multiple products on the market with a really

significant amount of historical data, both clinical and otherwise, and I would argue that you could use PK in that case to evaluate and produce a generic biopharmaceutical.

And I forget the third part.

DR. SIEGEL: I think it was a question about it being unethical to do certain trials. I won't comment on the pediatric part. I do want to make clear that my comment was not that it would be unethical to do these trials, but that there could be practical enrollment issues and sometimes ethical issues, depending on the setting, but certainly practical issues in asking people to participate in certain forms of--in certain trials, given that you would need to tell them that they have the option of taking the--not being in the trial and taking the innovator product, and that they would be randomized.

If the reason for doing the trial was to determine that the product wasn't inferior, then it might not excite a lot of people to be in such a trial, and, in certain settings, it raises ethical

issues, as it does even for innovators, where you have an effective treatment.

There are a lot of issues that arise, practical and ethical, in developing new treatments. There, in many cases, they hold out at least some hope of superiority.

As far as the PK/PD, I would say, as you heard from my comments, I think that clinical trials are, in virtually every case, going to be critically important.

I do think that PD markers, and Carole had a list of a number of things, like white cell count for CSF and hemoglobin and erythropoietin she spoke about, where I would agree with that point.

I wouldn't want to take a whole new class of drugs and look at its effect on hemoglobin or white count or glucose or whatever and say I know that that drug is effective in the treatment of whatever it is, anemia or neutropenia or diabetes.

But on the other hand, when you are within a well defined case or if you are, in this case, within even a high degree of similarity, I do think

that PD data will get you some important answers, but only to some of the questions, only to questions about some of the potential effects of the drugs.

DR. BEN-MAIMON: I would like to address the ethics issue. First of all, I think we did not say it was unethical to do clinical trials. What we said was it is unethical to do unnecessary clinical trials, and that is a very important discriminating factor.

When a trial provides important information to advance medical science, drug development, our understanding, quite honestly, I would say access to pharmaceutical products, it is absolutely ethical to do clinical trials.

We all deal with these issues of comparators and going into clinical trials. I mean, placebos are the utmost, when you've got a product that you know that works and you're looking at a placebo controlled trial.

So I think some of the issues that are raised are very important issues, but they are very

much--we really can deal with them and we have IRVs in place and things like that in order to help manage those risks.

DR. GERRARD: Terry Gerrard, TLG Consulting. I wanted to address Jay's concern about minor product differences and whether this poses an acceptable risk, because this seems to be maybe in conflict with common industry practice.

I know Jay was talking about a comparison between an innovator and a biogeneric, but innovators themselves often have minor product differences either due to manufacturing changes, lot-to-lot variation, and this is acceptable.

Sometimes preclinical studies are done, sometimes they're not. Sometimes clinical studies or PK studies. A lot of times they're not. And I think we, meaning industry and FDA, have accepted, for the past 10 to 15 years, that minor product differences are acceptable, that a product doesn't have to be identical. It has to be comparable, and we have addressed this through PK a lot of the times.

Now, the innovator can say we have a lot of the product history, that's true, but do they really have a full understanding of the clinical or

biological implications of the minor product differences? No, and that is probably unreasonable to ask of anybody.

So I think that we have accepted that minor product variances are an accepted risk if they have no effect on PK.

DR. SIEGEL: I think, in part, you are saying what I was trying to say, in part, which is that the question is not you do all the analytics and prove things to be identical, because that is just not likely to happen.

You do the analytics, you find minor differences, and you make determinations about whether those differences matter, and I think you said the same thing.

Now, in the case of an innovator product, you have a body of, at least some body of clinical data in which there is some variation in the product that was made and some information about

the effect of that variation on clinical outcomes.

So you have some range of clinical data to use to add to your determinations about which structural differences do and don't make a difference.

DR. GERRARD: But, Jay, we know we have marketed products out there now that are clearly different from the product that was used in phase three, and we really don't understand and probably are not expected to understand the implication of were those differences--well, we see efficacy, we continue to see efficacy of the marketed product, it's probably not substantial. We have accepted that.

DR. THOMAS: Adrian Thomas, Johnson & Johnson. I have a couple of comments. The first is I don't think it's a particularly strong argument to refer to products being marketed in other environments and say that they have been proven to be safe and efficacious without seeing the data around that, and I would be very interested to know what sort of products are

marketed, where, and what the nature of the data is supporting their safety and efficacy.

The second comment I would make is that one's perspective is clearly the shape of one's experiences. So when we talk about the bar not being different for generics as for innovators, I think whenever there is an established safety issue or concern that is not about a class or a product, it becomes incumbent upon those people marketing products or developing them to explore those safety issues fully.

It is not really adequate enough to say that 20 years ago you didn't have to do that, because 20 years ago, I think the science was very, very different.

DR. BEN-MAIMON: I totally agree with your second point. Clearly, if you take the small molecule world, we are held to standards identical to those today, not to when the innovator was actually developed or approved, and, clearly, standards change, GMPs change, specifications change, analytical methods change.

So I think we all agree that whatever biopharmaceuticals were being made, be they generic or brand, they would be developed to the standards

of the time with which they were being reviewed and approved.

From the standpoint of products marketed outside of the United States, I know we in the United States, although you speak with an accent, we in the United States like to believe that we are the only ones who can review and approve safe and effective products.

I beg to differ and I think as time goes forward, there are companies making these products in other parts of the world. As time goes forward, some of those products may ultimately come to the United States, and I hope that the agency will rigorously review that data and is sure that they are comparable, because they may be not comparable.

They may be, as I said, the "me too" type product, but, clearly, we can't ignore the fact that companies are manufacturing and producing insulin, GSCF, interferons, a whole host of

products, and patients are taking them, using them, and benefitting from them.

DR. THOMAS: I agree with you, but I think it is very important, when one makes an argument like that, to actually have some data around it so that we will understand what we are discussing, because simply referring to something in an abstract way conveys the message that perhaps isn't as scientifically underpinned as other messages.

DR. BEN-MAIMON: I would refer you to September's meetings. There was data presented by Sicor, if I remember correctly, on specific products, and it is all in the docket.

DR. NOVAK: Jean Novak, CBR International. Two-part question. That is, in the event that an abbreviated clinical path would be appropriate for either in toto biogenerics or a class of biogenerics defined, what would you envision the post-approval safety monitoring to look like and what would you think would be appropriate?

Secondarily, in light of--we've had a lot of discussion about immunogenicity and I think

there is going to be a lot more this afternoon, but with regard to immunogenicity as it pertains to efficacy, for example, a neutralizing antibody response that doesn't pose a safety issue, but, in fact, may result in increased dosing in the clinic, at will by the physician or even in a label recommendation, for example, I doubt it, but the point being how do you see these things, these differences that might occur later, again, after approval, when you don't necessarily have an active mechanism for following up on those kinds of issues?

DR. SIEGEL: As far as post-marketing, I would hope, if we embark on various potential policies that would bring new products to market, whether on the same standards or different standards, lowered or comparable efficacy testing, that we retain a system in which we can determine what product a patient is taking. I think that is critically important.

We learned in Johnson & Johnson, Centocor is a Johnson & Johnson company, from the PRCA

story, and we know from other events, too, that adverse events will emerge. They will emerge later and sometimes they will emerge as a result of changes in a product, sometimes just they're being used in a new setting.

We saw how difficult it can be to get at the case--that was an extensive investigation that took several years and over a $100 million to understand the cause of PRCA, but it was critically important because it was an important product, and critical to its importance was understanding what product each patient had received.

So I would simply say that as we move forward, we need to make sure that patients are traceable and that manufacturers have full responsibility and appropriate resources for ensuring that they can stand behind the products in the post-marketing setting.

DR. BEN-MAIMON: First of all, I hope we don't use a lower standard. I hope that we use a standard that ensures a comfort level that we are

all comfortable with.

But I would also say that you should recognize that the regulations for post-marketed products are identical for generics and for brands.

We do exactly the same pharmacovigilance, file exactly the same reports for our generic drug products as the brand industry does, and I would assume that that would be continued through as we moved into generic biopharmaceuticals.

Clearly, being able to find out who manufactured the product, trace the product, we've all talked about reimportation issues, it all relates to the same thing. It is essential.

Then, finally, I would say that for certain products, there may actually be additional requirements, depending upon the product and the scientific portfolio that is out there.

So we would look at a product by product basis and clearly we would be conscientious in making sure that we complied with all the regulations and all the post-marketing requirements and commitments.

DR. DI LIGERTI: Charlie Di Ligerti, Barr Labs. This question is for Dr. Siegel.

You underscored this morning the

difficulties in adequately powering clinical studies intending to detect low frequency events, adverse events.

When a brand product manufacturer makes a significant change in a marketed product, a change of the magnitude that necessitates a clinical study, what sort of equivalence criteria or acceptance criteria does that brand product manufacturer typically use to determine that the product pre-change and post-change are indeed equivalent with respect to low frequency safety events?

DR. SIEGEL: Those questions, of course, are going to be addressed on a product-by-product basis, but one of the critical factors is going to be how potentially important are uncommon adverse events.

Carole pointed out that in some settings, that the very low frequency is such that they are a

lower concern, but in some cases, of course, the nature of the event and the nature of the disease in the patient population, and PRCA would perhaps be a good example, are going to be more critical, I think you would find us and any manufacturer of EPO and any regulator of an EPO rather cautious about the type of data that would be necessary for a manufacturing change of an erythropoietin product.

Beyond that, I can't give numerical answers as to how you answer it. There's just a lot of factors that go into that.

DR. ORLOFF: Last question.

DR. STARK: Yafit Stark, TEVA Pharmaceutical. I have a question to Dr. Carole Ben-Maimon. In your talk, thanks for a very interesting talk, I would like to ask about a--you have described that in cases, that we need to run clinical studies. We like to see that the clinical studies will be targeted, meaning that they will ask a specific scientifically sound question.

I would like to ask you whether you feel that there should be some correlation with the

remaining uncertainty or the quality attributes to the clinical safety and efficacy while you are defining such a study.

DR. BEN-MAIMON: Yes. I think that was what was intended by my comment, which was that based on what our findings are, both from chemical and analytical testing, as well as our PK findings and potentially PD findings, you would design your studies or immunogenicity findings, you would design your clinical tests to answer specifically those questions rather than overall morbidity/mortality type answer or general efficacy type study.

DR. ORLOFF: Let me thank the speakers, again, Dr. Siegel, Dr. Ben-Maimon, and we will see you after lunch.

[Applause.]

[Whereupon, at 12:30 p.m., the session concluded.]

- - -